# Personal Photo Management and Preservation

**Chapter** · February 2018

DOI: 10.1007/978-3-319-73465-1_8



Andrea Ceroni
Forschungszentrum L3S

# Chapter 8
# Personal Photo Management and Preservation

Andrea Ceroni

**Abstract** Thanks to the spread of digital photography and available devices, taking photographs has become effortless and tolerated nearly everywhere. This makes people easily ending up with hundreds or thousands of photos, for example, when returning from a holiday trip or taking part in ceremonies, concerts, and other events. Furthermore, photos are also taken of more mundane motives, such as food and aspects of everyday life, further increasing the number of photos to be dealt with. The decreased prices of storage devices make dumping the whole set of photos common and affordable. However, this practice frequently makes the stored collections a kind of dark archives, which are rarely accessed and enjoyed again in the future. The big size of the collections makes revisiting them time demanding.

This suggests to identify, with the support of automated methods, the sets of most important photos within the whole collections and to invest some preservation effort for keeping them accessible over time. Evaluating the importance of photos to their owners is a complex process, which is often driven by personal attachment, memories behind the content and personal tastes that are difficult to capture automatically. Therefore, to better understand the selection process for photo preservation and future revisiting, the first part of this chapter presents a user study on a photo selection task where participants selected subsets of most important pictures from their own collections.

In the second part of this chapter, we present methods to automatically select important photos from personal collections, in light of the insights emerged from the user study. We model a notion of photo importance driven by user expectations, which represents what photos users perceive as important and would have selected. We present an expectation-oriented method for photo selection, where information at both photo- and collection-level is considered to predict the importance of photos.

---

Andrea Ceroni
L3S Research Center, Leibniz Universität Hannover, e-mail: ceroni@L3S.de

## 8.1 Introduction

Photos are excellent means for keeping and refreshing memories, they can illustrate situations we have gone through and serve as memory cues [290, 292] to bring back reminiscences of experiences, events, and people from our past. In the recent years we have been witnessing a huge increase in the production of photographs, mostly due to the wide spread of digital devices such as cameras, smartphones, tablets. People easily take hundreds or even thousands of photos during relatively short and memorable events, e.g., vacations, ceremonies, concerts, or depicting more mundane aspects of everyday life [308], like shopping, eating, working, free time. These numbers can amount to Terabytes of data over years. Even, considering only those uploaded in social media like Flickr, Facebook, Instagram and Snapchat, a study conducted few years ago [297] estimates that 500 million photos (most of them from personal collections) are uploaded to the Internet every day. In addition, this number is expected to double every year.

This scenario points out the significance of properly dealing with such increasing volumes of pictures. Due to decreasing storage prices and offers of cloud storage services, e.g., by Microsoft or Google, it is not a problem to store personal photos somewhere. As a matter of fact, directly dumping photo collections spending little or even no time in activities like pruning, editing, sorting, or naming has turned out to be a popular procedure [211]. This comes at a price: storage devices tend to become a kind of "dark archives" [228] of photo collections, which means that the stored pictures, although still available, are rarely accessed and revisited again in the future. The big size of the stored collections makes going through them such a tedious activity to prevent the viewers from accessing them at all. As an additional challenge, already discussed in Chapter 1, there is the risk of losing photos by a random form of "digital forgetting" [196]: over decades storage devices break down, and formats and storage media become obsolete, making random parts of photo collections inaccessible (Digital Obsolescence [374]). One example is how difficult it would be today to access photos stored years ago in .mos format in a floppy disk.

Both the threats of personal dark archives and digital forgetting raise the following question: how can photos be kept enjoyable and serve their original purpose as memory cues, where large photo collections tend to get dumped on hard disks and other types of storage? We propose a transition from dumped contents to more selective personal digital memories, supported by automatic methods for information value assessment, to support long-term personal data management. Regarding photo collections, this means identifying the most important photos from an entire collection and investing some effort to keep them accessible and enjoyable on the long run. Having a reduced sub set of important photos would make the revisiting easier and more pleasant for the user. However, understanding the importance of pictures to their owner for preservation and revisiting purposes is a complex process due to the presence of hidden factors, which are hard to model and capture automatically. These can be, for instance, memories, context, relationships to whom is in the picture, or simply personal tastes.

Therefore, the first part of this chapter summarizes a user study for a photo selection task where participants were asked to provide their personal photo collections and to select the subsets of photos that they would want to preserve and revisit again in the future. The study involved 35 participants, each one contributing at least one personal collection containing some hundreds of photos. The goal of this part is to better understand the human selection process for photo preservation and revisiting, identifying insights, patterns, and challenges that can shape the development of automatic selection approaches. Moreover, the gathered data will be employed for the development and evaluation of automatic selection methods. The user study was complemented by a survey, which we asked the participants to fill after completing the photo selection task.

In the second part of this chapter, we present and compare methods to automatically select important photos from personal collections for the sake of preservation and revisiting, inspired by the insights emerged from the user study. Many approaches to photo selection for summarization are centered around the concept of coverage, aiming at creating summaries that resemble the original collection as much as possible (see Section 8.2.2 for an overview). However, we believe that the complex decision making behind the selection of photos from personal collections, characterized by personal attachment due to memories, might reduce the importance of coverage. Therefore, we model a notion of photo importance driven by user expectations, which represents what photos users perceive as important and would have selected. We present an expectation-oriented selection method, where information at both photo- and collection-level (incorporating a relaxed notion of coverage) is considered to predict the importance of photos. We also investigate the role of coverage further by combining the expectation-oriented selection with an explicit modeling of coverage in different ways, showing that coverage plays only a secondary role in this task.

Our approach is an attempt to estimate Preservation Value (PV), previously introduced in Chapter 4, considering personal photos as specific information items. In fact, the notion of importance assigned to each photo by our model reflects what should be kept for future preservation, because the selection decisions collected during our user study and exploited for the development of our model have been taken exactly for that purpose. Furthermore, the input information extracted from photo collections covers different PV dimensions, compatibly with the characteristics of our scenario (e.g., the popularity dimension is not addressed as it refers to sharing and liking behaviors that are not very prominent in a personal scenario, where the data is rather kept private). The PV dimensions that we take into account are discussed in Section 8.4.2.1 and their roles within the selection process are summarized in Section 8.4.5.6.

The rest of the chapter is structured as follows. In Section 8.2, we outline related works and current approaches to photo selection. Section 8.3 describes the user study while the selection methods are presented and compared in Section 8.4. Finally, in Section 8.5 we summarize and conclude the chapter.

## 8.2 Related Work

The discussion of previous works relevant to the topic of personal photo selection and preservation is organized in two parts. The first one mentions previous empirical and rather qualitative user studies, while the second one reviews automatic approaches to the photo selection and summarization tasks.

### *8.2.1 User Studies and Surveys*

A considerable research effort has been dedicated to investigate issues related to photo management and preservation from a Human-Computer Interaction perspective [99, 211, 416, 426, 428]. Kirk et al. [211] introduced the notion of "photowork" as the set of activities performed with digital photos after capturing them and before any end usage like sharing or revisiting. One of their findings was that people spend little time in activities like reviewing, pruning, editing, sorting, because these are cumbersome and time consuming procedures. This fact clearly supports the topic and objective of this chapter. In the context of preservation of public photos, a qualitative study assessing their value for representing social history is reported in [99]. This study is mostly limited in (a) not considering personal photos and (b) the small number of photos considered. The evaluators were asked to rate five images, selected from Flickr, considering their worthiness for long-term preservation. Interestingly, the participants expressed a clear inclination to preserve all the pictures irrespective of their actual value. The authors hypothesized two possible reasons for this, namely the difficulty of anticipating a future information need and the effort required for organizing and pruning increasing amounts of data. In any case, they recognized this as a problem and pointed to the need of methodologies for information appraisal and selection.

Wolters et al. [426] investigated which photos from an event people tend to delete over time. In this study, described in Chapter 2, the participants took photos during a common event and then they were asked for deletion decisions at different points in time. While this work is certainly related to our study, which drew inspiration from it especially regarding the formulation of the survey, there are nevertheless some differences. Despite preservation ("keep") and "delete" decisions are related, we explicitly asked our evaluators to make selection decisions for the purposes of preservation and revisiting of images, rather than for deletion. Moreover, in our study the users were asked to make joint selection decisions (i.e., select a sub-collection) instead of making decisions for each individual picture in isolation. This is potentially a key difference, since selecting one photo might affect the decisions for other similar photos. Finally, instead of taking pictures of a common event explicitly for the study, we work with personal real-world collections belonging to diversified events. A subsequent work by Wolters et al. [428] presented a large-scale survey of 72 young people and students, with the goal of supporting the design of personal and mobile preservation systems. The main message of the study is coherent with what emerged from our survey: a large part of the participants acknowledge the importance of preserving photos for future generations. Interestingly however, only a small fraction of them carries out practices to support photo management and preservation. Another user study has been presented in [416], where participants wearing eye tracking devices were asked to select subsets of photos from two collections depicting two social events. This work focuses more on the selection process than on preservation matters. The survey on the aspects driving the selection process shares with our experiment both similarities (e.g., most of the highly rated aspects were subjective) and differences (e.g., quality was highly rated there).

### *8.2.2 Photo Selection and Summarization*

Automated photo selection has already been studied in various other contexts, such as, photo summarization [236, 365, 372, 397, 420], identification of appealing photos based on quality and aesthetics [235, 435], selection of representative photos [79, 416], and the creation of photo books from social media content [340]. We consider the task of selecting important photos from personal collections (e.g., for revisiting or preservation), which meet user expectations.

The work of Wang et al. [420] is probably the most related to ours, as their model of image importance does not explicitly include coverage and diversity aspects. They introduce the notion of "event–specific image importance", meaning that the importance of photos for selection purposes depends on the category of the event they belong to. The main assumption is that, within a photo collection depicting a certain type of event, the set of images commonly perceived as important by most people can be identified based on the event type. There are, however, substantial differences regarding the task definition and the way the ground truth was built. First, the ground truth was not gathered considering photo selection, since ratings were assigned by the evaluators to each image in isolation without explicitly deciding what subset of the collection should be kept. Second, individuals different than the collection owner rated the importance of images, potentially ignoring any personal attachment due to memories or hidden context. Image importance has been also considered in [235, 435], nevertheless, it was based on quality and aesthetic criteria. Instead, we explicitly consider selections preferences and expectations of users both for training our model and as evaluation criterion. Walber et al. [416] also consider human judgments to evaluate selections, but the users have to wear eye trackers when using the system to make automatic selections because gaze information is used as features in the model.

Different photo selection and summarization works consider coverage by identifying clusters of images based on time and visual content [79, 236, 340]. Differently, our approach does not impose such a strict notion of coverage but rather considers clusters and other global information together with image-level information, learning their different impact in a single model. The works in [306, 365, 372, 397] are closer to ours, as they consider coverage in a relaxed way as part of a multi-goal optimization, but they still consider coverage as a key component. Moreover, [365, 372] do not consider user assessments in their evaluation and make partial use of manually created text to associate semantic descriptors to images, while our method does not require any manual input, once the models for both feature extraction and importance estimation have been learned. Image collection summarization is performed in [306, 397] by applying structured prediction methods for learning weighted mixtures of submodular functions. The attention is drawn to two aspects that good summaries should exhibit, namely fidelity (coverage) and diversity, which are represented as a set of non-negative submodular functions and combined together in a single weighted submodular scoring function. There are two main differences with respect to the work presented in this chapter. First, their goal is purely summarization, aiming at optimizing coverage and diversity of output summaries, without considering whether they contain the most valuable pictures. This is strengthened by the utilization of the recall-based V-ROUGE metric (a criterion for summary evaluation inspired by the ROUGE metric [244], used for document summarization) within the loss function. Second, the way the ground truth has been collected is heavily oriented towards coverage: the evaluators, not the owners of the collections, were explicitly told to produce reference summaries that summarize the original collections in the best possible way, and those exhibiting low coverage were discarded. Conversely, we asked the collection's owners to select the most important photos according to their memories and perceptions, without any mention to coverage or diversity.

Besides [365, 372], other works in the literature rely on external knowledge to accomplish the task of image summarization [60, 353, 439]. Camargo et al. [60] combine textual and visual contents of a collection in the same latent semantic space, using Convex Non-Negative Matrix Factorization, to generate multimodal image collection summaries. Domain-specific ontologies are required as further input in [353]. They provide the knowledge about the concepts in a domain and are used to derive a set of ontology-based features for measuring the semantic similarity between images. Finally, [439] jointly leverages image content and associated tags and encodes the selection of images in two vectors, for the visual and textual domain respectively, whose non-zero elements represent the images to be included in the summary. The optimization process makes use of a similarity-inducing regularizer imposed on the two vectors to encourage the summary images to be representative in both visual and textual domains.

Summarizing, our approach is different from all the previous works under at least one of the following aspects: (a) our notion of photo importance is based on selection decisions made by people on their own photo collections; (b) we do not estimate photo importance using single indicators (e.g., quality, presence of faces, representativeness of the cluster a photo belongs to), but we rather learn the impact of such aspects through a single prediction model; (c) we use selection decision made by the collection owners themselves as ground truth for evaluation; (d) we do not rely on any kind of photo tagging or descriptive annotation provided manually.

## 8.3 User Study

As a preliminary step towards the development of automatic methods, we describe a user study conducted on a photo selection task, whose objective is the gathering of insights, challenges, and behaviors exhibited by humans when selecting personal photos for preservation and revisiting purposes [70]. Using their own photo collections depicting personal events, participants were asked to select a subset of photos that they would like to stay accessible and enjoyable in the future. Such data, i.e., the whole collections along with the selections done by the users, will be used for the training and evaluation of the selection methods described in Section 8.4. Upon completion of the task, the participants were also asked to fill a survey about it, which is described and analyzed in Section 8.3.2. This study is closely related to the one described in Chapter 2 and has been partially inspired by it. However, as already elaborated during the survey of the literature in Section 8.2.1, there are some important differences. First, we asked the participants to jointly select a sub-collection for the sake of preservation and revisiting instead of making “keep” or “delete” decisions for each image in isolation. Second, our study involves personal collections spontaneously taken and belonging to diversified events rather than photos of a common event taken explicitly for the study. This section is mainly organized into three parts: Section 8.3.1 elaborates on the setup of the study, Section 8.3.2 reports the insights learned from the study and the survey, in Section 8.3.3 we show a comparison between event-based clustering and human selections.

### *8.3.1 Task Setup*

The setup of the performed photo selection task involves the gathering of both participants and their photo collections, instructions on how the task should be accomplished, and, of course, the development of a software application to perform the selection in a comfortable way.

The experiment involved 35 users (28.6% females and 71.4% males) with 15 nationalities: 25.7% of the participants came from Greece, 17.1% from Germany, 11.4% from Italy, 11.4% from China, 5.7% from Vietnam, and the rest from Ethiopia, Turkey, Kosovo, Iran, UK, Thailand, Sweden, Brazil, Albania, and Georgia. Regarding their ages, 60.0% of the participants are between 20 and 30 years, 25.7% between 30 and 40, 11.4% between 40 and 50, 2.9% between 50 and 60.

Previous works mostly consider either public photo collections, for instance available on social media like Facebook and Flickr [60, 340, 353, 420], or pictures from a shared event in which all the evaluators took part [416]. One difficulty we see with using public collections of photos from different people, even if they attended the same event, is that according to the different experiences of the individuals in the event they might also have a different level of appreciation for the same photo, thus influencing their decisions. In contrast, we use personal photo collections. For instance, these can be photos from business trips, vacations, ceremonies, or other personal events the evaluator participated in. This means that each collection is not just a bunch of pictures, which might exhibit different degrees of quality and aesthetics, but there are experiences, sub-events, and memories that might influence the selection behavior. We decided to focus on such personal collections because we wanted to observe the personal photo selection decisions in a setting that is as realistic as possible. In total, 39 collections were used in the experiment (four users evaluated two collections), resulting in 8,528 photos. The size of the collections ranges between 100 and 625 pictures, with an average size of 219 and a standard deviation of 128.7. These collection sizes also emphasize the need for automated selection support, since manually browsing for photo selection becomes time-consuming. We asked users for further information about their collections, such as, the main topic of the collection, whether they were previously pruned (e.g., by discarding low quality photos), and when they were taken. Overall, 51% of the collections represent vacations, 30% business trips, and 19% other events like music festivals and graduation ceremonies. In addition, 23% of the collections were already pruned before the evaluation. The time when the collections were taken spans from 2007 to 2014 (64% in 2013-2014, 17% in 2011-2012, the rest in 2007-2010).

Since our task of selecting photos for preservation is not an everyday task for the users, it was important to find a good metaphor for supporting the task. After discussing a number of options with cognitive experts, we decided to use the metaphor of a “magic digital vault”, which incorporates the ideas of protection, durability, and a sort of advanced technologies to keep things accessible in the long-term. Therefore, the task consisted in selecting a subset of valuable photos to be put in the magic digital vault, which would protect them against loss and would ensure that they remain readable and accessible over the next decades.

To perform the photo selection task, we developed a desktop application, which enabled the participants to import their own collections and to select photos in a comfortable way. It is depicted in Figure 8.1, where the photos contained in the imported collection are displayed in the bottom panel, while the ones selected are shown in the top panel. Note that, faces appearing in Figure 8.1 have been blurred for the sake of privacy (only for inclusion in this book). The photos are selected and deselected by double-clicking on them, and they can be enlarged to inspect them better and appreciate their quality, although no explicit reference to the quality aspect was made in our instructions to the users. The photos in the collection were shown in the same order in which they were taken, since this makes the browsing, remembering, and selection easier and more realistic for the users. Nevertheless, we also made a preliminary evaluation where the photos were shuffled before being presented. This resulted in higher evaluation time and a higher mental effort for the selection process, because it made picking from a set of related photos very difficult. We verified that keeping the original order did not introduce any significant bias in the selection towards the early photos in the collection. This could have been a risk, since users might lose attention or even complete the selection without going through the entire collection.

Before starting the evaluation, the users were personally introduced to the photo selection task as well as to the application that they were asked to use. Further re-

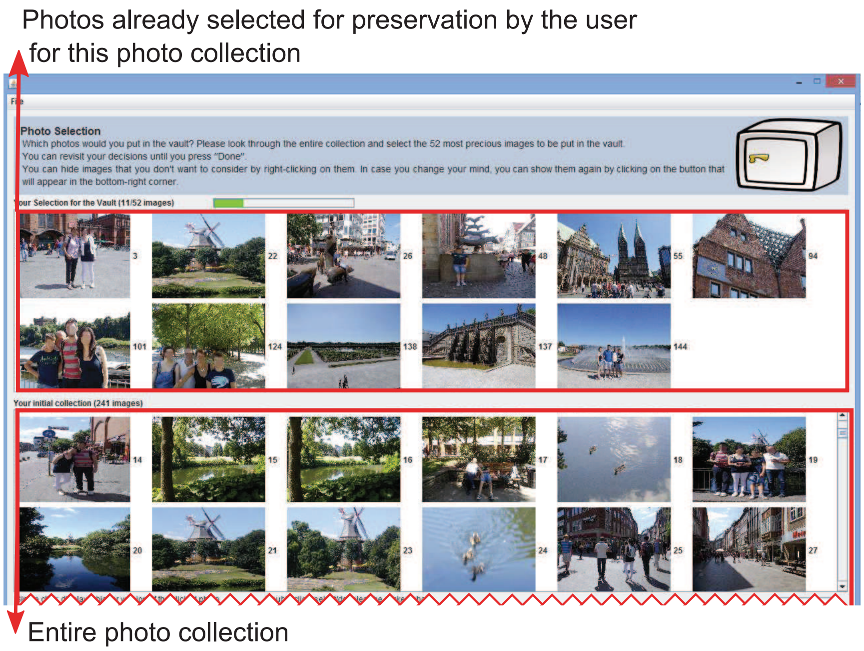


Fig. 8.1: GUI used by participants to browse the collections and select the photos to preserve.

marks and clarifications about both the task and the usage of the application were given, where needed. However, no guidelines were given about the criteria to use for selection, in order not to influence the selection process. After the users imported their collections, the application asked them to select 20% of photos from them for preservation and revisiting purposes. This selection percentage (20%) has been empirically identified as a reasonable amount of representative photos, after a discussion with a subset of users before the study. We also checked the adequacy of this chosen amount with the users in the survey by asking them whether they would have selected more photos if they could: 45% of them answered yes, the rest no. This balance means that 20% was a meaningful threshold, neither being to low (the majority of the users would have answered "yes" in this case) or too high.

### *8.3.2 Survey and Discussion*

After the photo selection step, the users were asked to fill a survey that can be conceptually split into two parts. The first group of questions refers to the scenario

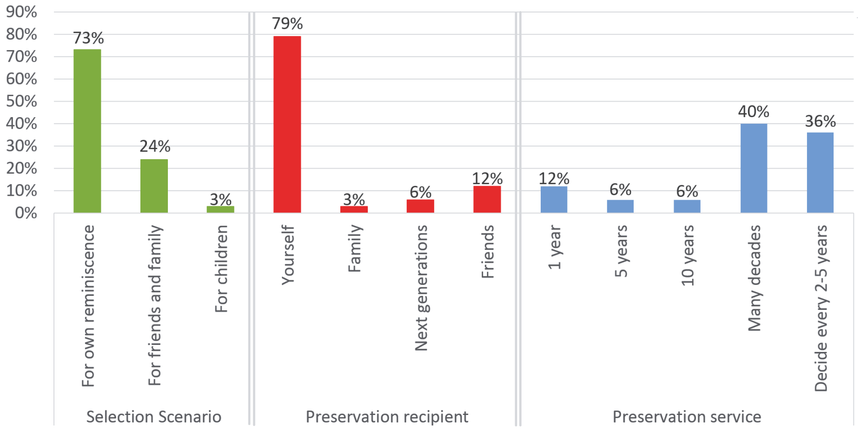


Fig. 8.2: Survey results with respect to preservation scenario, preservation target group, and preservation as a service.

of photo selection process for personal preservation, while the second one looks into the criteria that were considered during the selection.

Regarding the first group of questions, the users were asked to provide information about (a) which scenario they had in mind when selecting the photos; (b) for whom they are preserving the photos; (c) whether they would be ready to pay, and for how many years, if preservation was a paid service. The answers to each question were posed as multiple choices and are reported in Figure 8.2. The answers to questions (a) and (b) reveal that the long-term preservation process is centered around the owner of the photos: more than 70% of the evaluators said that they thought about own future reminiscence when they selected the photos, and almost 80% indicated themselves as a main consumer of the preservation outcome. Looking at the preservation as a valuable service to be paid (question (c)), the evaluators were mostly split into two groups: either being ready to pay for many decades (39%) or needing flexibility to make new preservation decisions every 2-5 years (36%). In both cases, these answers highlight a clear need for preservation of personal photo collections.

In the second group of questions, we suggested different photo selection criteria and asked the users to rate how much each criterion was considered during the selection. The suggested criteria, which are in line with the insights on "keep" and "delete" decisions in [426], were rated via star ratings on a scale between 1 and 5 (5 stars mean very important, 1 means not important at all). The criteria along with statistics about their ratings are reported as box plots in Figure 8.3. Note, that medians are represented as horizontal bold bars, while sample mean is indicated with a bold cross. For the sake of clarity, we grouped the criteria into three classes: "content-based criteria" refer to objective and subjective measures for individual photos such as quality, typicality (i.e., how suitable it is for serving as an iconic

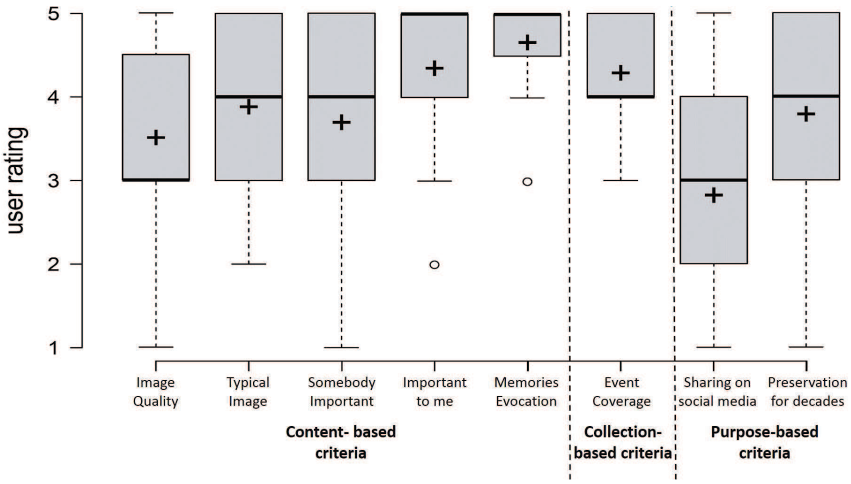


Fig. 8.3: Boxplots of the different selection criteria.

summary of the event), the presence of important people in them, whether photos are generally important, and the evocation of memories; “collection-based criteria” - here represented by coverage of events - consider a photo in the context of its collection; “purpose-based criteria”, indicating the importance of different selection goals (in our case, sharing and preservation).

An important finding of this evaluation is that the objective quality of photos is rated as the second least important selection criterion, after the sharing intent. This shows that quality and aesthetics, although being important and used for general-purpose photo selection [235], are not considered very important in case of selecting photos for preservation. In contrast, criteria more related to reminiscence, such as event coverage, typical image, and “the picture evokes (positive) memories” are all rated high, with highest ratings for memory evocation. The remaining two criteria “picture is important to me” and picture “shows somebody important” refer to the personal relationship to the picture and are also both rated high. These results anticipate that the task of predicting photos to be selected for long-term preservation is likely to be difficult, since many of the criteria that are rated high, e.g., memory evocation, personal importance and “typical image”, are difficult to assess for a machine, because they contain a high level of subjectivity. Another complicating fact is that there is no single dominant selection criterion, but a combination of highly rated criteria. In these ratings, we can observe differences with respect to the ones given to the partially overlapping set of criteria reported in [416], where photos on shared events were used and the selection was not directly related to preservation and reminiscence. In that work, much higher ratings are given to criteria such as quality, whereas event coverage and importance of depicted persons are rated relatively low (although with high variance). Interestingly, photos that capture a memory are also rated high in this case.

### *8.3.3 Image Clustering and Human Selections*

We analyze the applicability of current selection and summarization approaches to the scenario of personal photo selection for preservation, highlighting possible issues that they might face in this situation. The main uncertainties in applying state-of-the-art methods to our task are (a) that they are developed with other photo selection scenarios in mind and (b) that they often do not compare the performances of their output with selections done by users. They, for example, identify sub-sets of photos that provide comprehensive summaries of the initial collections [340, 372, 397], without checking if the summary meets the user expectations, or they consider judgments based on more objective criteria such as aesthetics [235, 435]. Since a wide part of the state-of-the-art methods for photo selection and summarization considers clustering and/or coverage for generating selections and summaries (as discussed in Section 8.2.2), we clustered photos by applying the event-based clustering technique described in Chapter 3 (Section 3.5) and compared the clustering results with the human selections. This analysis is corroborated by the fact that the event coverage criterion, representable through clustering, has been identified as important during our study (Section 8.3.2).

In our opinion, one of the main risks of applying clustering to emulate human selections for long-term preservation is that not all the clusters might be important for the users. There might be photos from a sub-event that the user either simply does not like or considers less important than others. We supported this hypothesis by counting the number of human-selected photos in each cluster identified in our collections. As to be expected, only for a few clusters (7.3%) all the photos within the same cluster were selected. However, for a considerable part of the clusters (43%) no photos were selected at all. Given these statistics, the selection done by any pure coverage–based method that picks an equal number of photos from each cluster will contain at least 43% of them that would not have been selected by the user. Another statistics worth to be mentioned refers to the possibility for cluster–based selections of picking centroids as representative photos. From our collections, it resulted that only the 26% of the centroids was actually selected by the users. This reveals that information about how much a photo is representative of a wider group is only one of the aspects considered by the users when selecting photos.

Finally, making the assumption that bigger clusters might be more important for the users (as indicated by the users' choice to take more photos that capture that part of the event), we consider the size of the clusters with respect to the number of user-selected photos that they contain. Figure 8.4 shows the correlation between relative size of clusters ($x$ axis) and the percentage of selected photos in them ($y$ axis). It is possible to observe that the selections done by the users result in many clusters with few selected photos in each, which is coherent with the notion of coverage. However,

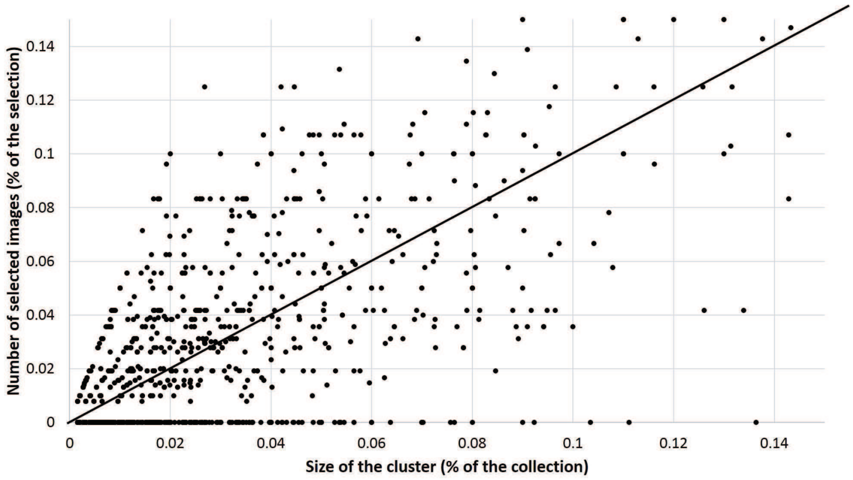


Fig. 8.4: Amount of selected photos in clusters (with respect to the size of selection) versus relative size of clusters.

what is more interesting is that the size of the cluster seems to be only marginally correlated with the importance of the cluster (i.e., the number of selected photos it contains). This is potentially another limitation for all those methods that select an amount of photos from each cluster proportionally to its size.

## 8.4 Photo Selection

We present in this section an automatic method [71] to identify, within big personal collections, those photos that are most important to the user, in order to invest more effort for keeping them accessible and enjoyable in the future. The availability of such a method alleviates the problems of "digital forgetting" and "dark archives", discussed in Section 8.1, which affect the archival of images and their access, respectively. From one side, preservation effort could be invested only on those photos that are worth to be preserved for the owner. From the other side, having a reduced sub set of important photos would make the revisiting and enjoying easier and pleasant for the user. Moreover, to foster adoption, such automated selection method has to keep the level of user investment low. We do not rely on any additional user investment such as photo annotation with text [340, 365, 372] or eye tracking information [416], because we believe it is exactly the reluctance of further investment that lets large photo collections unattended on our hard disks. To alleviate errors in automatically generated selections as well as accommodate user preferences, our approach can be regarded as a semi-automatized procedure, where the user can interact with it and modify the suggested selections.

When developing methods for semi-automatic photo selection, it is important to consider human expectations and practices. Photo selection is a complex and partially subjective process, where the selection decision taken for a given image both affects the decisions for other photos and depends on the ones already selected. For this reason, many state-of-the-art methods for photo selection and summarization are driven by the aspect of coverage, which means attempting to create summaries that resemble the content of the original collection as much as possible. Some of them perform a two-step process of first clustering the photo collection (for reflecting sub-events in the collection) and subsequently picking the most representative photos from the clusters [236, 340]. Others [365, 372, 397] consider coverage as part of a multi-goal optimization, along with the concepts of quality and diversity within the summary. While coverage surely plays an important role for many photo selection tasks (see e.g., [416]), we believe that the complex decision making behind the selection of photos from personal collections, characterized by subjectivity and personal attachment possibly due to memories, might reduce the importance of coverage. For instance, considering photos taken during a trip, the user might want to discard the ones depicting boring or joyless moments.

Therefore, we model a multifaceted notion of photo importance driven by user expectations, which represents what photos users perceive as important and would have selected. User expectations have been acquired during the study described in Section 8.3, where participants have been asked to provide their own photo collections and to select those most important to them for preservation and revisiting. We present an expectation-oriented method for photo selection, where information at both photo- and collection-level is considered to predict the importance of photos (Section 8.4.2). This information consists of: (a) concept detection, to capture the semantic content of images beyond aesthetic and quality indicators; (b) face detection, reflecting the importance of the presence of people within photos; (c) near-duplicate detection, to take the redundancy of many pictures of the same scene as a signal of importance, and to eliminate very similar ones; (d) quality assessment, since good quality photos might be preferred in case of comparable photos. This is complemented by (e) temporal event clustering and, more generally, collection-level information, to reflect the role of coverage in photo selection. The impact of the different features is learned through a single model to predict the importance of each photo. Information regarding the selections performed by the users from their own collections is explicitly used to train the selection model, so that the predicted importance of photos represents what the user would have selected. For sake of comparison, in Section 8.4.3 we investigate how the expectation-oriented selection can be combined with more explicit ways of modeling coverage, showing that coverage plays only a secondary role in this task.

Before delving into the details of the selection method, a general consideration on the comparison between the features considered in the model and the user study presented in Section 8.3 has to be done. The aspects that resulted to be important from the user study, e.g., evocation of positive memories, image typicality, personal importance of photos, are highly subjective and not directly recognizable by a machine, especially when only relying on the visual content without any other contextual information. Given these challenges and constraints imposed by the task, our attempt to address the insights emerged from the study is threefold: (a) we model event coverage, which resulted to be an important aspect in the user study, through clustering and the hybrid selection methods described in Section 8.4.3; (b) we employ concept detection to model more semantic and abstract aspects; (c) we also include image quality, although perceived as not very important within the user study, for the sake of comparison with the other features.

### 8.4.1 Overview

The problem that we tackle in this chapter can be formalized as follows.

**Definition 1** *Let a photo collection P be a set of N photos, where* $P = \{p_1, p_2, \ldots, p_N\}$. *The photo selection problem is to select a subset S of size* $\theta$ *(*$S \subset P$ *and* $|S| = \theta$*), which is as close as possible to the subset S* that the user would select as the photos most important to her, i.e., S meets user expectations.*

We represent each photo collection as a set $C = \{P, CL, ND\}$, where $P$ is the set of original photos, and $CL$ and $ND$ are sets of clusters and near-duplicate photos identified in the collection, respectively. A cluster $cl \in CL$ contains a set of photos $P_{cl}$ grouped together with respect to a defined notion of similarity, whereas a near-duplicate set $nd \in ND$ is a set of highly similar photos $P_{nd}$. Each photo $p \in P$ is modeled as a set of features $p = \{\mathbf{q}, \mathbf{c}, F, t\}$, where $\mathbf{q} \in \mathbb{R}^{n_q}$ is the quality vector of the photo, $\mathbf{c} \in \mathbb{R}^{n_c}$ is the concept vector of the photo, $F$ is the set of faces $f$ appearing in the photo, $t$ is its timestamp. Each face $f = \{f_l, f_s\}$ is described by its location $f_l$ and relative size $f_s$ in the photo. For each photo $p$, we will estimate the importance value $I$ using the extracted features.

Figure 8.5 depicts the overview of our approach to photo selection. Given a photo collection, we extract information from the photos it contains by applying different image processing techniques described in Chapter 3 and in [71], such as concept detection, image quality assessment, face detection, event clustering, near-duplicate detection. Our main approach is named expectation-oriented selection (Section 8.4.2), which learns to generate selections by taking into account user selections from personal collections as training data. Furthermore, we present three different hybrid selection methods (coverage-driven, filtered expectation-oriented, optimization-driven), with the goal of investigating whether our method can be improved by combining it with state-of-the-art methods that explicitly consider coverage. The hybrid selection methods will be discussed in detail in Section 8.4.3.

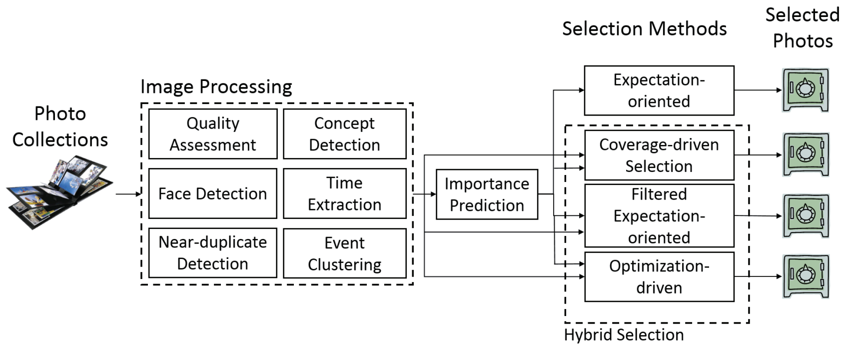


Fig. 8.5: Approach overview of automatic photo selection.

### *8.4.2 Expectation-oriented Selection*

The photo selection model presented in this section aims at meeting human expectations when selecting photos that are most important to the user from a collection, for revisiting or preservation purposes. We believe that selecting photos that are important to a user from personal collections is a different task than generating comprehensive summaries: the set of images important to the user might not be a proportioned subsample of the original collection. For this reason, we do not impose a strict notion of coverage but rather consider clusters and other global information as a set of features, along with photo-level features, learning their different impact in a single selection model by mean of supervised Machine Learning.

While we do not employ Deep Learning techniques, they can be used for either computing new features or replacing the computation of the current ones, still leaving the rest of the approach intact. For instance, image representations learned by Convolutional Neural Networks (e.g. GoogLeNet [5]) can be used for Concept Detection as shown in Section 3.2. This is indeed one of our goals for the next future.

A key characteristic of our features is that they do not require any manual annotation (e.g., tags, textual descriptions, file names) or external knowledge, differently from other works [340, 353, 365, 372] that make partial use of manually created text associated to photos. This means that the user does not have to invest time and effort in preparing the photos before feeding them into our system.

#### 8.4.2.1 Features

Four groups of features have been designed to be used in the photo selection task, based on the information extracted from photos as presented in Chapter 3 and [71].

The correspondence between these features and the PV dimensions defined in Chapter 4 is made explicit in their descriptions.

### Quality-based Features

They consist of the 5 quality measures, namely blur, contrast, darkness, noise, and their fused value (weighted pooling using Minkowski metric), which have been extracted as in [71], following the procedure presented in [281]. They are all numeric features whose values are between 0 and 1, where 0 represents the best quality and 1 the worst. The assumption behind using this information is that users might tend to select good quality photos, although their impact seems to be less important in subjective selections of humans as emerged from previous work [416] and from our user study (Section 8.3). Nevertheless, quality might probably play a role in case of near-duplicate images with different quality: the user would pick the best one in these cases. This family of features corresponds to the quality PV dimension defined in Chapter 4.

### Face-based Features

The presence and position of faces in a picture might be an indicator of importance and might influence the selection. Some people might prefer photos depicting persons instead of landscapes, others might like group photos more than single portraits. We capture this by considering, for each photo, the number of faces within it as well as their positions and relative sizes. Faces have been detected through the approach presented in [71], which combines several face detectors (all incorporating the Haar-like-feature-based detector by Viola & Jones [413]) to maximize the number of detected faces. Then, each photo is divided in nine quadrants and the number of faces and their size in each quadrant are computed, resulting in 19 features: two for number and size of faces in each quadrant, plus an aggregated one representing the total number of faces in the photo. These features, to some extent, can be associated to the social graph PV dimension defined in Chapter 4, because the presence of people in a picture could indicate relationships between the appearing people and the owner of the photo. The notion of who is in a picture, for instance obtainable via face clustering and tagging, would provide more precise and complex relationships among the owners and the people appearing in their collections. However, this would also introduce additional effort for the user, who would have to manually make the system aware of the kind of relationship with respect to any unknown person found in a new collection.

### Concept-based Features

High-level and semantic information has been thoroughly investigated in the past years within the scope of digital summarization (e.g., [365, 372]). The semantic content of photos, which we model in terms of concepts appearing in them, is expected to be a better indicator than low-level visual features, because it is closer to what a picture encapsulates. We consider the 346 concepts defined as part of the TRECVID 2013 benchmarking activity [7] and previously mentioned in Chapter 3, Section 3.2.3.1. The concept set includes both abstract concepts, such as "joy", "entertainment", "greeting", and more concrete concepts, such as "animal", "building", "mountain". We trained a Support Vector Machine (SVM) as concept detector for each of them, using the TRECVID 2013 dataset (described in Chapter 3, Section 3.2.3.1) as training corpus. We used SIFT, SURF, and ORB local descriptors and their color variants [6] for visual feature extraction. Then, PCA was applied on each descriptor for reducing their dimensionality to 80 and VLAD encoding [20] was applied to calculate the final image representation. The applied methodology is described in [6] in more detail. Having such detectors available, we associate to each photo a vector of 346 elements, one for each concept, where the *i*-th value represents the probability for the *i*-th concept to appear in the photo. The correspondence between this class of features and the PV dimensions is not strict and depends on which concepts are included in the concept space. Concepts might be related to the gravity dimension, in case they represent aspects related to the events in the collection, or to the social graph dimension, in case they represent appearance of people, groups, or crowds.

### Collection-based Features

All the previously mentioned features are extracted from photos in isolation. However, when users have to identify a subset of important photos, instead of just making decisions for each photo separately, the characteristics of the collection a photo belongs to might influence the overall selection of the subset. For the same reasons, but moving to a finer granularity, it might be worth considering information about the cluster a photo belongs to. This family of features is a representative of the coverage PV dimension. For each photo, we consider collection-based features to describe the collection and, if any, the cluster and near-duplicate set the photo belongs to. Regarding the whole collection, we consider its size, the number of clusters and near-duplicate sets in the collection, the number of not near-duplicate photos, the size of the clusters (avg, std, min, max) in the collection, the size of near-duplicate sets (avg, std, min, max) in the collection, the quality of the collection (avg, std), the number of faces in the collection (avg, std, max, min). Regarding clusters, we first perform event–based clustering by applying the method described in Chapter 3, Section 3.5. Then, given the cluster a given photo belongs to, we compute its size, its quality (avg, std, max, min), and the number of faces within it (avg). Finally, since the redundancy introduced by shooting many pictures of the same scene

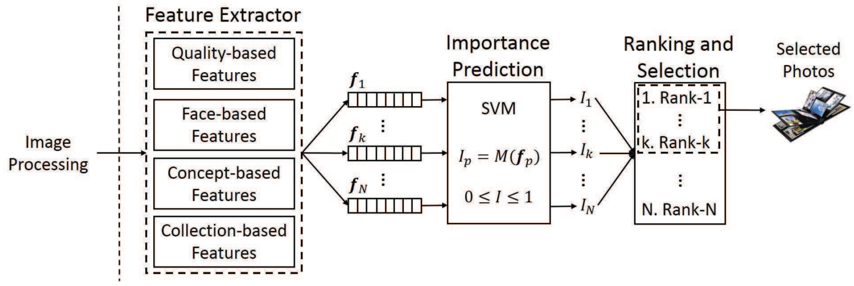


Fig. 8.6: Workflow of the importance prediction and photo selection.

can be evidence of its importance for the user, we also extract features regarding whether the given photo has near-duplicates or not, as well as how many they are. Near–duplicates are detected by mean of the methodology described in Section 3.4. Shooting many similar photos of the same scene can be regarded as a form of "investment", because the user puts effort in replicating a scene to ensure its availability and quality.

#### 8.4.2.2 Importance Prediction and Ranking

Given a set of photos $p_i$, their vectors $\mathbf{f}_{p_i}$ containing the features presented above, and their selection labels $l_{p_i}$ (i.e., "selected" or "not selected") available for training, a prediction model represented by an SVM is trained to predict the selection probabilities of new unseen photos, i.e., their importance (see Section 8.4.4 for details regarding the training process). Figure 8.6 shows how the importance prediction and ranking of photos is performed for new unseen collections. First, feature vectors $\mathbf{f}_p$ are constructed based on the information extracted from the collections as described before and the importance of each unseen photo $p$ is computed as:

$$I_p = M(\mathbf{f}_p) \tag{8.1}$$

which is the probability of the photo to be selected by the user. Second, once the importance of each photo in the collection has been predicted, they are ranked based on these values and the top-$k$ are finally selected. The parameter $k$ represents the requested size of the selection and has to be specified in advance. The choice of $k$ will be discussed during our evaluation (Section 8.4.4).

## 8.4.3 Hybrid Selection

In Section 8.4.2 we have presented an expectation–oriented photo selection model to explicitly meet selection decisions of humans via Machine Learning. As the evaluation will show (Section 8.4.5), our expectation-oriented selection clearly outperforms state-of-the-art methods for photo selection based on explicit modeling of coverage. However, given the wide exploitation of the concept of coverage in many state-of-the-art methods, we want to better understand its role in photo selection, in order to see if and in which way our method can be improved by combining it with explicit consideration of coverage. The notion of coverage resulted to be highly important from our user study (Section 8.3) as well, which is another motivation for further investigating its potential contributions and limitations. It is interesting to note that, despite the participants declared coverage as highly important, the selections that they made in the study exhibited a poor degree of coverage.

We propose and investigate three ways of combining our importance prediction model with coverage-oriented photo selection methods, denoted "hybrid selection" methods and described hereafter. The coverage PV dimension, although kept into account within the expectation-oriented selection via the collection-based features (Section 8.4.2.1), becomes more dominant in this new family of selection methods. While our discussion is centered around the role of coverage, it is worth mentioning that also the diversity PV dimension is considered within one of the hybrid methods (described in Section 8.4.3.3).

### 8.4.3.1 Coverage–driven Selection

The coverage-driven selection is based on the widely used two-step process of first clustering and subsequently picking photos from the clusters. First, for a given collection $C$, a set of clusters $CL_C$ is computed as described in Chapter 3 (Section 3.5) and the importance $I(p)$ of each photo $p \in P_C$ is computed according to our importance prediction model (Equation 8.1). Given the clusters $CL_C$, we use the importance $I(p)$ for each photo $p \in P_C$ to pick an equal number of top-ranked photos from each cluster in order to produce the selection $S$ of required size $k$.

#### Cluster Visiting

When picking photos from each cluster, there are different possible ways of iterating over them until the requested size of the selection is reached. After experimenting a number of alternatives, we identified a round-robin strategy with a greedy selection at each round as the best performing one. The pseudo-code is listed in Algorithm 2. Given an initial set of candidate clusters $CL_{cand}$, the greedy strategy in each step selects the cluster $cl^*$ containing the photo $p^*$ with the highest importance, according to the prediction model $M$. The photo $p^*$ is added to the selection $S$ and removed from its cluster $cl^*$. The cluster $cl^*$ is then removed from the set of candidate clus-

**Algorithm 2** Coverage–driven Selection (Greedy)

```
Input: clusters CL, size k, prediction model M
Output: selection S

Set S = ∅
while |S| < k do
    Set CL_cand = CL
    while |CL_cand| > 0 do
        {cl*, p*} = get_most_important_cluster (CL_cand, M)
        S = S ∪ {p*}
        P_cl* = P_cl* − {p*}
        CL_cand = CL_cand − {cl*}
        if |cl*| = 0 then
            CL = CL − {cl*}
        end if
    end while
end while

return S
```

ters for this iteration, and the greedy strategy is repeated until the candidate set is empty. Once it is, all the not empty clusters are considered available again and a new iteration of the cluster visiting starts. This procedure continues until the requested selection size $k$ is reached. We also experimented with a regression model to predict the number of photos to select from each cluster, but it did not lead to satisfactory results.

### Cluster Filtering

The intuition behind cluster filtering is that not all the clusters identified in a collection are equally important to the user. For instance, considering photos taken during a trip, there might be some of them depicting exciting moments along with other more boring situations, which the user might want to discard. We tackle this issue by proposing a cluster filtering method to automatically predict the clusters that are not important for the user, in order to ignore them when picking photos from each cluster. We train a classifier (SVM) to detect and filter out clusters which are not important to the user. First, each cluster is described with the following features: size, quality vector (avg, std), average concept vector, number of faces (avg, std, min, max), number of near-duplicate sets and near-duplicate photos in it, near-duplicate sets size (avg, std, min, max), photo time (avg, std, min, max), photo importance (avg, std, min, max). The label associated to a cluster is "good" if it contains at least one selected photo, "bad" otherwise. Given a training set made of clusters $c_i$, their corresponding feature vectors $\mathbf{f}_{c_i}$, and their classes $l_{c_i}$, an SVM is trained and the learned model $N$ is used to predict the class $L = N(\mathbf{f}_{c_{new}})$ of new unseen clusters $c_{new}$. Details regarding the training process are reported in Section 8.4.4.

Given the clusters $CL_C$ in a collection and a classifier trained on a different portion of the dataset, applying cluster filtering removes from $CL_C$ all those clusters that are classified as *bad* by the classifier. The iteration and picking phase are then performed only with the remaining "good" clusters.

#### 8.4.3.2 Filtered Expectation–oriented Selection

The coverage-driven selection is characterized by two steps: first clusters are identified and handled by possibly filtering and sorting them, and then photos in each cluster are ranked based on their predicted importance. Differently, within the filtered expectation-oriented selection, we give priority to importance prediction. The photos in a collection are first ranked based on the predicted importance and then cluster filtering is applied. The result is a ranked list of photos, where those belonging to clusters classified as "bad" have been removed. Note that the second phase of this paradigm, which contains cluster filtering in our case, can incorporate any other computation that exploits cluster information. The way photos are selected after applying cluster filtering is the same as the one described in Section 8.4.2.2: the selection $S$ of size $k$ is created by choosing the top–$k$ photos in the list.

#### 8.4.3.3 Optimization–driven Selection

Besides applying clustering, another way of explicitly incorporating coverage into a photo selection process is to consider it as part of a multi-goal optimization problem. This has been done in [372] to generate representative summaries from personal photo collections, with the objective of having concise sub-collections that resemble the original one as much as possible. In more detail, in this work quality, coverage, and diversity are jointly optimized and the optimal summary $S^*$ of a requested size $k$ is defined as:

$$S^* = \underset{S \subset P_C}{\arg\max}\, F\left(Qual\left(S\right), Div\left(S\right), Cov\left(S, P_C\right)\right) \tag{8.2}$$

where $Qual\left(S\right)$ determines the interestingness of the summary $S$ and it aggregates the "interest" values of the individual photos within $S$, $Div\left(S\right)$ is an aggregated measure of the diversity of the summary measured as $Div\left(S\right) = \min_{p_i, p_j \in S, i \neq j} Dist\left(p_i, p_j\right)$, and $Cov\left(S, P_C\right)$ denotes the number of photos in the original collection $C$ that are represented by the photos in the summary $S$ with respect to a concept space.

We incorporate our expectation-oriented selection within this framework, creating the optimization–driven selection, by computing the $Qual\left(\cdot\right)$ function in Equation 8.2 based on the importance prediction model (Equation 8.1), that is:

$$Qual\left(S\right) = \sum_{p \in S} M\left(p\right) \tag{8.3}$$

Since part of the concepts in [372] are discrete categorical attributes, associated to photos using textual information and external knowledge bases not available in our task, we binarized the elements of our automatically detected concept vector (which includes the probability that a given concept appears in the photo) by using a threshold $\tau$ such that $c_i = 1 \; if \; c_i > \tau$, and $c_i = 0$ otherwise. The threshold has been empirically identified as $\tau = 0.4$ as the value that led to the most meaningful binary results. The rest of the calculation of the $Div(\cdot)$ and $Cov(\cdot)$ functions in Equation 8.2 is performed as in the original work. In more detail, the distance between two photos, used to measure the diversity within a summary, is computed based on exif features, time, and concept vectors (as in the original work, however we use the automatically extracted concepts), while the coverage of a summary is calculated based on the number of photos in the original collection that are represented by the ones within the summary in a concept space (considering binarized concepts vectors when needed).

Regarding the resolution of Equation 8.2, which is an NP–Hard problem, we experimented with the different approaches presented in [372] and the best performing one consisted in combining quality, diversity, and coverage in a linear way:

$$S^* = \underset{S \subset P_C}{\arg\max} \left[ \alpha \cdot Qual(S) + \beta \cdot Div(S) + \gamma \cdot Cov(S, P_C) \right] \tag{8.4}$$

and performing a greedy optimization, which has proved performance guarantees (please refer to [372] for further details). We will discuss the values used for the $\alpha, \beta, \gamma$ parameters in the experimental analysis.

## *8.4.4 Experimental Setup*

### 8.4.4.1 Dataset

For our experiments we use personal photo collections with importance judgments given by the owners of the collections as dataset. These can be photos from business trips, vacations, ceremonies, or other personal events a person participated in. We decided to focus on personal collections because we wanted to observe the personal photo selection decisions in a setting that is as realistic as possible. This gives us a ground truth for assessing user expectations.

Given the unavailability of such a dataset of real-word personal collections, with selections done by the owners based on their perceived importance, we considered the data collected during the user study previously described in Section 8.3. As a short reminder, participants were asked to provide their personal photo collections and to select the 20% that they perceive as the most important for revisiting or preservation purposes. The selection percentage (20%) was empirically identified as a reasonable amount of representative photos, after discussing this matter with a subset of participants before the study. In order to make the evaluation results more statistically significant, we expanded the originally collected dataset (Section 8.3.1) by repeating the same evaluation procedure with other participants and photo collections. Such extended dataset consists in 18,147 photos organized in 91 collections and belonging to 42 users. The collection sizes range between 100 and 625 photos, with an average of 199.4 (SD = 101.03).

Near-duplicates have been detected and filtered by considering the centroid of each set as representative photo, as done in [79]. For sets containing two photos, the one with better quality is chosen as representative. Similarly to [360], each representative is marked as selected if at least one photo in its set has been marked as selected, and marked as not selected otherwise.

#### 8.4.4.2 Evaluation Metrics

Since the overall goal of our work is emulating the human behaviors in selecting the subsets of photos from a personal collection, we compare the automatic selections generated by our methods with the ones done by the users.

The selection methods presented in this chapter can generate a selection $S$ of size $k$ from the original collection, where $k$ can assume different values. We evaluate the different methods considering the precision P@$k$ of the selection $S$ of size $k$ that they produce, computed as the ratio between number of photos in $S$ that were originally selected by the user and the size of $S$. Since the collections in our dataset have high size variability (from 100 to 625 photos), absolute values of $k$, although traditionally used in Information Retrieval tasks, would result in selecting very different relative portions of the collections depending on their sizes. This makes the impact of the selection different among collections. We, therefore, decided to express $k$ as a percentage of the collection size, instead of an absolute value. In particular, we compute the precision for $k = 5\%, 10\%, 15\%, 20\%$, which are indicated as P@5%, P@10%, P@15%, P@20%, respectively. We concentrate the discussion on P@20%, because our ground truth was gathered by asking users to select the 20% of their collections. We will also give comments about the recall of the selections generated by the different methods.

The 91 collections in our dataset have been split by 10-fold cross validation (used for training and evaluating the classifiers) and all the values reported in the rest of this section are averaged over the test sets of each split. Statistical significance tests were performed using a two-tailed paired t-test and significant improvements are marked as $^{\blacktriangle}$ and $^{\vartriangle}$ (with $p < 0.01$ and $p < 0.05$, respectively). If not stated otherwise, the significance outcome reported in the tables always refers to the comparisons with both the baselines described in Section 8.4.4.4.

#### 8.4.4.3 Parameter Settings

The classifiers employed in this chapter for importance prediction and cluster filtering, built using the SVM implementation of LibSVM[1], have Radial Basis Functions as Kernels and their hyper-parameters are the following. The ones of the SVM used within the expectation-oriented method for importance prediction are $C = 1.0$, $\gamma = 1.0$, while the SVM used for cluster filtering has parameters $C = 3.0$, $\gamma = 1.5$. All of them were tuned by grid search and 10-fold cross validation.

#### 8.4.4.4 Baselines

We compare our method with two baselines, one based on clustering and one representing the optimization framework presented in [372].

Clustering

Similarly to what was described at the beginning of Section 8.4.3.1, for a given collection $C$, a set of clusters $CL_C$ is computed. The selection is built by iterating the clusters, temporally sorted, in a round–robin fashion and picking at each round the most important photo from the current cluster (until the requested selection size is reached). Instead of using our expectation-based model, the importance of each photo $p \in P_C$ is modeled as

$$I(p) = \alpha \cdot \left\| \mathbf{q}_p \right\| + (1 - \alpha) \cdot \dim(F_p) \tag{8.5}$$

which is a weighted sum of the quality vector of the photo and the number of faces in it. This notion of image importance covers different works in the literature, for instance [236, 340]. We experimented with different values of the parameter $\alpha$, identifying the best value as $\alpha = 0.3$, which gives more importance to the number of faces in the photos. We report the performances obtained with this parameter value in our evaluation.

Summary Optimization

We implemented the approach presented in [372] as another baseline, where summaries are generated by optimizing quality, coverage, and diversity as in Equation 8.2. It differs from the hybrid method described in Section 8.4.3.3 in how photo importance is modeled, as here the expectation-oriented model is not considered. Instead, the quality of summaries is computed by summing the "interest" of photos in it, defined as a measure dependent on photo quality and presence of portraits, groups, and panoramas. We computed the interest of photos as in the original work,

[1] http://www.csie.ntu.edu.tw/~cjlin/libsvm/

using the concepts “face”, “3 or more people”, and “landscape” available in our concept set to represent portraits, groups, and panoramas, respectively. Also the diversity and coverage of summaries are computed coherently with their original computation, as already described in 8.4.3.3. Assigning equal weights to the $\alpha, \beta, \gamma$ parameters gave us the best results, thus we will report the performances for only this setup in the following evaluation, denoting it *SummOpt*.

## *8.4.5 Results*

The discussion of the results is organized as follows. First, we show the performances of our expectation-oriented selection with respect to the baselines, discussing the impact of different features subsets in the selection (Section 8.4.5.1). We also analyze the correlation between single features and selections in Section 8.4.5.2. Second, we present the results of the hybrid selection methods and we compare them both with the baselines (Section 8.4.5.3) and with the expectation-oriented selection (Section 8.4.5.4). Third, we make a general comparison of the methods based on recall performances (Section 8.4.5.5).

Besides providing the numeric performances, we discuss and compare the dominant criteria behind each method, and we map such analysis to the PV dimensions introduced in Chapter 4. These considerations are summarized in Section 8.4.5.6.

### 8.4.5.1 Expectation–oriented Selection

This section presents the evaluation of our expectation-oriented selection with respect to the two baselines defined in Section 8.4.4.4. Different importance prediction models have been trained by using the subsets of the features described in Section 8.4.2.1, so that the impact of different groups of features on the precision can be analyzed. Since each group is linked to part of the PV dimensions (Chapter 4), our analysis provides insights about the importance of the dimensions in the context of personal photo selection for preservation.

To reduce the dimensionality of the *concepts* features, consisting of 346 numerical values, we performed feature selection based on Information Gain [145] and we kept the top–160 features for training. The amount of 160 features has been empirically identified as a compromise between simplicity and expressiveness of the model. We did not apply any feature selection on *quality* and *faces* features because their dimensionality is small.

The results for different selection sizes ($k$) are listed in Table 8.1. The two baselines exhibit comparable performances, with *SummOpt* performing slightly better for all considered values of $k$ (5%, 10%, 15%, 20%). Regarding our model, *quality* features are the ones that perform weakest individually, which has already been observed for other photo selection tasks [416]. This corroborates the idea that low quality photos might be kept anyway because they contain and recall memories and

Table 8.1: Precision of the expectation-oriented selection, distinguishing different sets of features.

| | P@5% | P@10% | P@15% | P@20% |
|---|---|---|---|---|
| *Baselines* | | | | |
| Clustering | 0.3741 | 0.3600 | 0.3436 | 0.3358 |
| SummOpt | 0.3858 | 0.3843 | 0.3687 | 0.3478 |
| *Expectation-oriented Selection* | | | | |
| quality | 0.3431 | 0.3261 | 0.3204 | 0.3168 |
| faces | 0.4506$^\blacktriangle$ | 0.3968$^\blacktriangle$ | 0.3836$^\vartriangle$ | 0.3747$^\vartriangle$ |
| concepts | 0.5464$^\blacktriangle$ | 0.4599$^\blacktriangle$ | 0.4257$^\blacktriangle$ | 0.4117$^\blacktriangle$ |
| photo-level | 0.5482$^\blacktriangle$ | 0.4760$^\blacktriangle$ | 0.4434$^\blacktriangle$ | 0.4266$^\blacktriangle$ |
| **all (Expo)** | **0.7124**$^\blacktriangle$ | **0.5500**$^\blacktriangle$ | **0.4895**$^\blacktriangle$ | **0.4652**$^\blacktriangle$ |

events important to the user. *Faces* features alone already show better performances than the baselines: the presence, number, and position of people in photos, largely used as one selection criterion in other works, is indeed a meaningful indicator of importance. The performance achieved when only using *concepts* features is better than the ones of *quality* and *faces*: they are able to capture the semantic content of the photos, going beyond their superficial aesthetic and quality. Examples of concepts with a high importance in the model are “person”, “joy”, “entertainment”, and “crowd”. The model trained with the combination of all the aforementioned features, denoted *photo-level* because the features are extracted from each picture in isolation, slightly improves the performance of using concept features alone. This indicates that leveraging quality and faces features in addition to semantic measures, such as concepts, can ameliorate the overall performance.

If we include global features for each photo representing information about the collection, the cluster, and the near–duplicate set the photo belongs to, we get a comprehensive set of features, which we call *all*. Similarly to the case of *concepts* features, we performed feature selection based on Information Gain on the whole set of *all* features and we retained the top–200 features for training. Again, the set size has been empirically identified as a compromise between simplicity and expressiveness of the model. The precision of the selection for this global model further increases for every selection size: this suggests that decisions for single photos are not taken in isolation but they are also driven by considering general characteristics of the collection the photo belongs to: e.g., number of photos, clusters, average quality of photos in the collection and in the same cluster, how many duplicates for the photo there are. This is a point of distinction with respect to state-of-the-art methods (represented by the two baselines), because our selection approach does not strictly handle collection-level information by imposing clustering (*Clustering*) or optimizing measures like coverage and diversity along with photo importance only based on quality and presence of people (*SummOpt*). It rather takes this global information in consideration in a flexible way through a set of features, whose impact to the

Table 8.2: Top-30 features ranked by Information Gain with respect to the class.

| Info Gain | Feature Name | Info Gain | Feature Name |
|---|---|---|---|
| 0.10836 | ND of photos | 0.01561 | Avg aggr. quality in collection |
| 0.02569 | Images without ND in collection | 0.01538 | Std ND set size |
| 0.02258 | Min darkness in cluster † | 0.01523 | Min ND set size |
| 0.02251 | Std aggr. quality in collection | 0.01469 | Std faces in collection |
| 0.02240 | Norm of concepts in collection | 0.01440 | Concept "person" |
| 0.02189 | Count of faces in photo | 0.01414 | Count of faces in cluster† |
| 0.02177 | Avg size of ND sets in collection | 0.01321 | Std aggr. quality in cluster† |
| 0.02144 | Avg contrast in cluster† | 0.01306 | Concept "dresses" |
| 0.02009 | Max cluster size in collection | 0.01291 | Concept "joy" |
| 0.01863 | Avg contrast in collection | 0.01273 | Avg blur in cluster† |
| 0.01760 | Count of central faces in photo | 0.01147 | Avg blur in collection |
| 0.01732 | Avg count of faces in collection | 0.00952 | Concept "two people" |
| 0.01610 | Min clusters size | 0.00889 | Concept "entertainment" |
| 0.01609 | ND sets in collection | 0.00873 | Contrast of photo |
| 0.01565 | Size of central faces in photo | 0.00826 | Concept "girl" |

selection is learned from user selections and expectations. The expectation-oriented model using all the available features (named *Expo* in the rest of the evaluation) leads to a relative improvement of 38.5% and 33.75% over *Clustering* and *SummOpt* respectively, considering P@20%, and even higher improvements when considering smaller values of $k$ (90.4% and 84.6% for P@5%).

The P@20% metric is of primary importance because we asked users to select exactly 20% of their collections during the data acquisition. However, another point of discussion is the trend of precision performances over different values of $k$: all the models reach higher precision values for smaller selection sizes. This can be due to the presence of a limited number of selected photos that are relatively easy to identify for the methods, which give them highest selection probability.

Summarizing, modeling different promising aspects in terms of features and flexibly combining them through Machine Learning leads (except when using quality information alone) to consistent and statistically significant improvements over state-of-the-art summarization and selection methods.

#### 8.4.5.2 Feature Analysis

For sake of completeness, in Table 8.2 we report the top–30 features ranked based on the Information Gain with respect to the class (i.e., user selections). Despite the presence of similar and redundant features, the table still provides an overview of the features that are correlated to the class the most. The symbol † for features related to clusters means that the cluster containing the input photo is considered. For instance, given an input photo, the feature "Min darkness in cluster" represents the minimum darkness over all the images within the cluster the input photo belongs to. The first-ranked feature, whose Information Gain value is significantly higher than the ones of the other features, represents the number of near-duplicates that the input photo has. This reveals that the redundancy introduced by taking many shoots of the same scene is a strong signal of importance for that scene. Besides this feature, the other ones in the table have much smaller and similar Information Gain values. Many other high-ranked features are computed considering global information from clusters and collections, which confirms what already discussed: the decisions taken for single photos implicitly take into account general characteristics of the collection the photo belongs to. Features computed based on faces are also important, namely the total number of faces in the picture and the number and size of faces in the center of the photo. Quality is mostly considered in relation to collections and clusters (i.e., quality statistics with respect to the whole collection or a given cluster). A relatively low number of features represent concepts, which is somewhat counter intuitive if compared with the selection results of the *concepts* features reported in Table 8.1. Nevertheless, the high performance values, if compared to those of *quality* and *faces* features, might be due to the combination of many concept features, although they are not all top-ranked.

#### 8.4.5.3 Hybrid Selection

This section discusses the precision of the hybrid selection methods presented in Section 8.4.3 with respect to the baselines, along with a comparative analysis between the different hybrid selection methods. The results are listed in Table 8.3, where they have been split based on the three different classes of hybrid selection described in Section 8.4.3. For coverage-driven selection, we report results of different combinations: *basic* refers to the coverage–driven selection which only uses our importance prediction model defined in Section 8.4.2.2 as photo importance measure, picking photos in a round-robin fashion from clusters temporally ordered; the term *filtered* means the use of cluster filtering, while the presence of the term *greedy* indicates the use of the greedy visiting strategy. The filtered expectation-oriented selection is denoted *F-Expo*.

For the optimization-driven method, we experimented with the different optimization methods described in [372] after introducing our importance prediction model in place of the original importance measure used in that work ($Qual(\cdot)$). We found out that the best performing method was still the greedy optimization of a linear cost functional combining importance, diversity, and coverage (Equation 8.4) but with a parameter combination that gives more importance to the quality of the photos (0.6 *Qual*, 0.3 *Cov*, 0.1 *Div*). We consider the results of this setup in the following evaluation. This difference in weights with respect to the *SummOpt* baseline already anticipates that our expectation-based measure of importance has a bigger impact in the performances than the native quality measure defined in [372]. The method will be referred to as *SummOpt*++.

The results in Table 8.3 show that all hybrid methods outperform the baselines, with statistical significance, showing that the inclusion of the importance prediction model to assess photo importance has a strong impact compared to the baselines

Table 8.3: Precision of the hybrid selection methods.

| | **P@5%** | **P@10%** | **P@15%** | **P@20%** |
|---|---|---|---|---|
| *Baselines* | | | | |
| Clustering | 0.3741 | 0.3600 | 0.3436 | 0.3358 |
| SummOpt | 0.3858 | 0.3843 | 0.3687 | 0.3478 |
| *Coverage-driven Selection* | | | | |
| basic | 0.4732$^{\blacktriangle}$ | 0.4113$^{\blacktriangle}$ | 0.3902$^{\vartriangle}$ | 0.3809$^{\vartriangle}$ |
| filtered | 0.5351$^{\blacktriangle}$ | 0.4617$^{\blacktriangle}$ | 0.4325$^{\blacktriangle}$ | 0.4170$^{\blacktriangle}$ |
| filtered+greedy | 0.6271$^{\blacktriangle}$ | 0.4835$^{\blacktriangle}$ | 0.4391$^{\blacktriangle}$ | 0.4262$^{\blacktriangle}$ |
| F-Expo | 0.7065$^{\blacktriangle}$ | 0.5502$^{\blacktriangle}$ | 0.4863$^{\blacktriangle}$ | 0.4600$^{\blacktriangle}$ |
| SummOpt++ | 0.7115$^{\blacktriangle}$ | **0.5533**$^{\blacktriangle}$ | **0.4937**$^{\blacktriangle}$ | **0.4708**$^{\blacktriangle}$ |
| Expo | **0.7124**$^{\blacktriangle}$ | 0.5500$^{\blacktriangle}$ | 0.4895$^{\blacktriangle}$ | 0.4652$^{\blacktriangle}$ |
| *Filtering with Oracle* | | | | |
| greedy+oracle | 0.6499$^{\blacktriangle}$ | 0.5107$^{\blacktriangle}$ | 0.4665$^{\blacktriangle}$ | 0.4484$^{\blacktriangle}$ |
| F-Expo+oracle | 0.7150$^{\blacktriangle}$ | 0.5606$^{\blacktriangle}$ | 0.4982$^{\blacktriangle}$ | 0.4753$^{\blacktriangle}$ |

methods, which model photo importance with simple functions of quality and people occurrence. Similarly to the performances of the expectation-oriented models, both the absolute precision values and the improvements with respect to the baselines increase for decreasing $k$.

Focusing on the *coverage-driven selection*, the results in Table 8.3 also show that cluster filtering increments the precision of the *basic* approach of an amount between 9.48% (P@20%) and 13.1% (P@10%). The greedy visiting strategy leads to improvements as well. Statistical significance tests revealed that the improvements introduced by *filtered* and *filtered+greedy* are statistically significant.

Comparing the results of the different hybrid selection methods, *F-Expo* and *SummOpt++* achieve better precision performances than the coverage-driven methods, and a t-test confirms that these improvements are statistically significant. This shows that the measure of photo importance modeled by our importance prediction has a bigger impact in the precision of the selection than coverage, and those methods that strictly model it through clustering (*coverage-driven selection*) get a smaller benefit when incorporating the expectation-oriented model. On the other side, methods that either give priority to expectations (*F-Expo*) or consider expectations, coverage, and global information in a flexible way via optimization (*SummOpt++*) can better exploit the expectation-oriented model.

#### 8.4.5.4 Expectation vs. Hybrid Analysis

In this section we make a comparative analysis between the expectation-oriented selection model exploiting all the available features (*Expo*), and the hybrid selection models. Considering Table 8.3, we can observe that the performances of *Expo* are better or comparable with the ones of the hybrid-selection models. In particular, the improvements of *Expo* with respect to the *coverage-driven* methods are statistically significant. The only improvements over *Expo* (which anyway are not statistically significant) are obtained when considering methods that prioritize expectations (*F-Expo*) or possess a relaxed consideration of coverage and global information in general (*SummOpt++*). These results further support our assumption that, for the photo selection task involving personal data, a strong consideration of coverage overstresses this aspect as a selection criterion. Instead, the users might not follow a strict idea of coverage when making selections, generating selections that are not as proportioned samples of the original collections as purely coverage-based methods would suggest. Only for the methods with a more flexible consideration of coverage the performances are similar to the pure expectation-oriented method.

Cluster filtering is an attempt to eliminate clusters uninteresting to the user, and in order to further alleviate this aspect we conducted experiments considering only important clusters, i.e., those ones containing at least one selected photo. This is done by assuming to have a perfect classifier, i.e., an "oracle", to filter out not important clusters and to focus the hybrid selection strategies only on the important ones. Although getting improvements compared to *filtered+greedy* and *F-Expo*, the performances when using such oracle, reported in the bottom part of Table 8.3, did not lead to consistent and statistically significant improvements with respect to *Expo*. *Greedy+oracle* does not beat *Expo*, while *F-Expo+oracle* only introduces a limited and not statistically significant improvement. These results show that the aspect that mostly drives user selections and expectations is the personal perception of importance, although this can produce unbalanced selections which are not representative of the original collection. Another problem related to clustering, even considering the important ones, might be the decision of how many photos to pick from each of them.

#### 8.4.5.5 Recall-based Analysis

We make a comparative analysis of the different method based on recall. The motivation of considering recall is that a user might accept to increase the size of the automatically created selection in order to include more important photos than the ones included when remaining strict to the ideal size of 20% (considered during the user study). In Table 8.4 we show the recalls of the best performing methods from each selection class, computed for different selection sizes. Note that R@20% always coincides with P@20%, since users were asked to select 20% of their collections. The results are coherent with the analysis already done for the precision: both the expectation-based model and the hybrid-selection methods outperform the baselines, and the former is overall better then or comparable to the latter class. Only methods that prioritize user selections (*filtered expectation-based*) or consider expectations, coverage, and global information in a flexible way via optimization (*optimization-driven selection*) can reach slightly higher recall values than the one of the expectation-based model. In the future, this consideration could be the start-

Table 8.4: Recall of different selection methods.

| | R@20% | R@30% | R@50% | R@75% |
|---|---|---|---|---|
| Clustering | 0.3358 | 0.4555 | 0.7000 | 0.9231 |
| SummOpt | 0.3478 | 0.4354 | 0.6884 | 0.9253 |
| Expo | 0.4652$^{\blacktriangle}$ | 0.5310$^{\blacktriangle}$ | 0.7356$^{\vartriangle}$ | 0.9310 |
| filtered+greedy | 0.4262$^{\blacktriangle}$ | 0.5129$^{\vartriangle}$ | 0.7232 | 0.9231 |
| Filtered Expo | 0.4600$^{\blacktriangle}$ | 0.5361$^{\blacktriangle}$ | 0.7433$^{\vartriangle}$ | 0.9275 |
| **SummOpt++** | **0.4708**$^{\blacktriangle}$ | **0.5408**$^{\blacktriangle}$ | **0.7405**$^{\vartriangle}$ | **0.9315** |

ing point for a photo selection method that maximizes recall, or at least considers it in the learning model along with precision-based criteria.

#### 8.4.5.6 Features and Preservation Value Dimensions

We now summarize the main insights obtained from this work, linking the results of the photo selection methods to the high-level dimensions of the PV introduced in Chapter 4.

From the results reported in Section 8.4.5.1, *quality* features are the ones that perform worst, revealing that the quality PV dimension is not of primary importance for preservation in personal scenarios. As an example, one might want to keep a photo because it evokes memories of the time when we took the photo, despite its low quality. The *faces* class of features alone, although performing better than *quality* features, was not as a good indicator as expected from its common usage in photo summarization and selection. The introduction of face clustering and tagging, to know who is actually appearing in the photos and what the relationship with the collection owner is, might probably help make the social graph dimension more important. However, this would also require a certain investment of the user in tagging and annotating, as well as the awareness of social relationships, which are both not assumed to be available in the considered scenario.

Since a wide part of the state-of-the-art methods for photo selection and summarization [79, 340, 365, 372] considers clustering and, more generally, coverage as primary criterion for generating selections and summaries, we applied selection methods based on temporal clustering and on summary optimization to our scenario. These high expectations on the coverage dimension were not confirmed by the experimental results. We observed that emphasizing coverage, either strictly by selecting photos fairly from each cluster or more flexibly via summary optimization, did not yield significant improvements over the pure expectation-oriented selection, which incorporates global information in a more relaxed way through a set of features. The only positive result related to coverage is the high correlation between the presence of near-duplicates and selection decisions (Table 8.2), which shows that people tend to shoot many similar photos of what they like the most and is most important to them. However, this fact is more related to the concepts of redundancy and investment than coverage. In our opinion, one of the main pitfalls of stressing coverage to emulate human selections from personal collections for preservation is that not all the clusters are usually equally important for the users. There might be photos from a sub-event that the user either simply does not like or considers less important than others. The optimal parameter values identified for the optimization-driven selection (Section 8.4.3.3), jointly considering importance, coverage, and diversity, showed that also the diversity dimension had a low impact in the selection. While being widely considered for photo summarization, diversity resulted to have only a marginal role in emulating user selections for preservation.

### *8.4.6 Personalization*

Although the expectation-oriented method presented in Section 8.4.2 has been proved to be more effective in meeting user expectations than state-of-the-art approaches based on coverage, it applies the learned selection model for any user and collection. Nevertheless, the photo selection process (especially for personal data) can be highly subjective and the factors that drive the selection can vary from individual to individual [321, 360, 435]. General selection models, although capable of representing common selection patterns (e.g., photos depicting people might be usually appreciated), might be improved by considering the preferences of each single user separately and derive personalized models for them. Some users might be particularly interested in photos depicting many people, while others might prefer pictures with landscapes or buildings. Besides variations in the set of appreciated concepts, also selection aspects that are ignored by some people might become more important for others. It is therefore worth spending some thoughts on how personalized selection models that adapt to the preferences of different users could be developed. To this aim, we have performed a preliminary study in [147] to investigate how personalized models can be derived from the photo selection approach described in Section 8.4.2, denoted "general model" hereafter. We highlight the applied methodology and its findings in the rest of this section.

#### 8.4.6.1 Methodology

A recurrent matter in Machine Learning is continuously managing new data, so that the existing model can be updated to accommodate new information and to adapt to it. Two common approaches for updating the model to new incoming data are Online Learning [47], where the model is updated only considering the new data, and Incremental Learning [66], where the model update considers the old training data along with the incoming data. We considered the latter strategy and re-train the model each time new data (i.e., selection decisions) was provided by the user because, in our scenario, the updated model has to be aware of the entire data available,

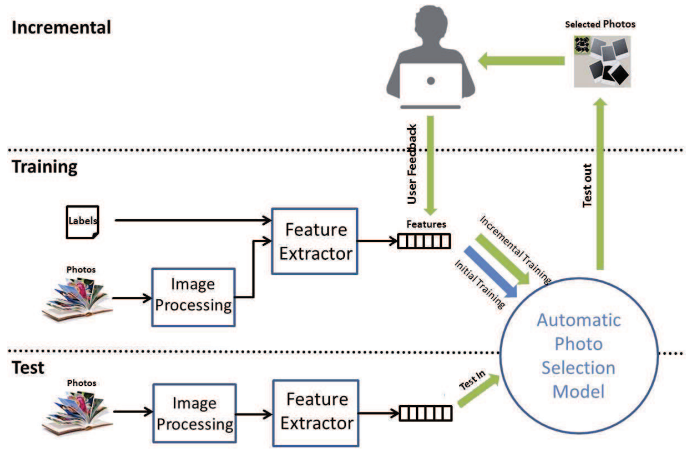


Fig. 8.7: Overview of the personalization of the general photo selection model.

not just of the most recent one. Although efficient and effective incremental versions of off-line learning algorithms exist (e.g., [66]), we performed the model update by including the new data in the training set and re-train the model from scratch. We implemented such more straightforward but functionally equivalent approach because our scenario does not impose strict temporal constraints for the model update, thus making the efficiency benefit of incremental versions of secondary importance. The time taken by a user to produce a new collection (e.g., after a trip or vacation) can be considered sufficient to re-train the model with the whole available data. Should the temporal constraints of the envisioned scenario become stricter, the incremental version of the employed algorithm could be plugged in without changing the functionalities of the whole application.

The personalization workflow is summarized in Figure 8.7, which emulates the application of the personalized model in real-world settings. The personalized photo selection models, one for each given user, are built by re-training the model every time that a new collection is imported and the automatic selection done by the current selection model is revised by the user. The annotated photo collections available to train the general model are first pre-processed through image processing techniques and features are extracted from them, in the same way as described before in this chapter. For each new collection provided by the user, a first selection is made by the trained general model as described in Section 8.4.2.2 and the selected photos are displayed to the user, who gives feedback revising the automatically generated selection. The training dataset is then expanded by adding the feedback data and the general model is retrained with the updated training dataset. Iterating this process, it is expected that the gap between user expectations and model's selections gets lower, due to the adaptation of the model towards the selection preferences of the user. This workflow represents the envisioned behavior once the whole system has been finalized and released to the end user. However, in order to easily repeat evaluations when designing and implementing the model, we collected the data from each user once for all, i.e., users evaluated all the collections from scratch without revising any automatically generated selection. Although we are aware that the selections done by the user starting from an automatically generated selection might differ from those done when selecting photos from scratch, repeating the evaluation multiple times when designing the system would have been unfeasible for the users. Moreover, acquiring evaluations done from scratch is unbiased from the initial selection proposed automatically.

Usually, the adaptation of a system within the initial rounds of user interactions is affected by the so called "cold-start problem": there is not enough (or even not at all) training data to let the model adapt to the user. This holds in our scenario as well, where the selection model might not make proper predictions due to the lack of annotated collections in the initial training set. We considered two ways of building the initial training set. One consists in using one annotated collection of the given user as initial training set. The other is based on using annotated collections from other users to train the initial selection model, hopefully boosting the adaptation of the model to a given user when a limited amount of personal training data is available. The latter approach is based on the assumption that, despite the subjectivity of the task, common selection patterns exist and could be captured through a sample of selections done by other users.

#### 8.4.6.2 Findings

We used the dataset already described in Section 8.4.4 as basis of our experiments. In order to assess personalization performances, we consider users who contributed at least 5 collections as test users. Among the overall 91 photo collections, there are 11 users who provided at least 5 collections (10 users contributed 5 collections, 1 user contributed 6 collections) which result in 56 collections totally. Afterward, the original dataset is split into two parts: one part contains 35 collections from 31 users, whereby each user provided at most 2 collections, which is named *general dataset*; another part contains 56 collections from 11 users, whereby each user provided at least 5 collections, which is called *personalized dataset*.

Given the aforementioned general distinction between *general* and *personalized* dataset, we evaluate the performances of the model update over different rounds of adaptation. The *personalized dataset* is split based on users where each one owns 5 collections (one user owns 6). At each iteration $k$, for each user with $N$ collections, $k$ collections are added to the initial training set to learn the personalized model of the user, and $N-k$ collections are used for testing. We considered three ways of building training sets. In the *stand-alone* procedure the initial model is trained with one random collection of the user, and the model update is incrementally done considering the remaining collections of the same user. In this case we are considering each test user in isolation, ignoring any data from others. The *collaborative* strategy fills the initial training set with all the collections within the *general dataset*. This case represents the situation where, in absence of large amount of annotated personal data for training, annotated collections of other users are used to alleviate the cold-start problem. The *user–agnostic* method, similarly to the *collaborative* case, uses the *general dataset* as initial training set. However, at each iteration, instead of including collections of the user under consideration, we add randomly selected collections from the other test users. This case is motivated by the assumption that, if one collection, which is not from the user that we are considering, is included in the training set at each iteration, then the adaptation performances should be smaller than including collections that are from the user that we are considering. This would highlight the importance of incorporating selection information of the user in the training set when making selections for new collections of the same user.

We observed from the experiments that the precision of both *stand-alone* and *collaborative* increases at each iteration, i.e., with the increase of the number of user's collections considered for training the model. This suggests that having a selection model partially aware of the user preferences (by exploiting a certain amount of the selection behavior in the training phase) can improve the precision of new unseen collections of the same user. The precision of *collaborative* was higher than the one of *stand-alone*, especially at the first iterations, showing that the selection data from other users can alleviate the cold-start problem. We also measured the relative gain obtained by each strategy between any two consecutive iterations. The gain of *stand-alone* at each iteration resulted to be higher than the one of *collaborative*, because the initial model was weaker (due to the limited training set) and the inclusion of new training collections had a higher impact on the learning. Comparing *user-agnostic* and *collaborative*, the former exhibited an almost null or even negative gain over iterations, while the latter led to a bigger and increasing performance gain at every iteration. This demonstrated that the increase of performance at each iteration was due to the inclusion of a new collection of the same user in the training set and not simply caused by expanding the training set at each iteration, since in this case the gain of *user-agnostic* should have been higher as well.

As a conclusion, this evaluation led to promising results, showing that (a) including new annotated collections for the same user when training the model can benefit the selections on new unseen collections of the same user, and (b) exploiting annotated collections from other users as initial training data can boost the system performances in cold-start scenarios. It is important to clarify that the standard deviation observed in these experiments was relatively high. This can be due to a mixture of aspects, such as a limited size of test set (both in terms of users and iterations) and intrinsic changes of difficulty among collections of the same user. For this reason, although a promising user adaptation emerged from this study, the inclusion of an extended amount of users, collections, and iterations would help make the results more evident and statistically significant.

## 8.5 Conclusions

In this chapter, we considered the problem of keeping personal photo collections enjoyable over time. Given the explosion in the production of digital photos within the recent years and the common practice of merely dumping such data on cheap storage devices or using storage services, the stored photo collections are rarely accessed and revisited afterward. To some extent, their content tends to be forgotten because the big collection size makes their revisiting a fatiguing process.

As a remedy, we proposed a selective approach to long-term data management that aims at identifying what is most important to the user and investing in the longevity and enrichment of this content, in order to make the future revisiting more enjoyable and less tedious. The development of such automated method was preceded by a user study, to lay the foundations of the task and better understand its challenges.

The user study was centered around a photo selection task where 35 participants contributed their own collections, from which they selected the photos most important to them, namely the ones that they would like to preserve for future revisiting. One important outcome was that many hidden and subjective criteria (memory evocation, personal importance, photo typicality) were rated high, anticipating the difficulty of automatizing the selection task. Moreover, the more objective criterion of photo quality was rated as less important. Another aspect emerged as important was coverage, which means that the set of selected pictures should fairly represent the content of the original collection. Although this was stated by the participants, their selections exhibited a poor degree of coverage.

Afterward, we presented an expectation-oriented method for photo selection exploiting an extensive set of photo- and collection-level features, to estimate the long-term photo importance based on user expectations. The evidence of user expectations has been derived from the personal data provided during the user study and has been used to train the selection model. The goal of this method is supporting users in selecting the most important photos for creating an enjoyable sub-collection of a personal collection for preservation and revisiting purposes. Since a wide part of state-of-the-art methods is driven by the concept of coverage, which resulted to be highly rated in our user study as well, we also investigated how to combine the expectation-oriented selection with more explicit modelings of coverage. Experiments with real world photo collections showed that (a) our method outperforms such state-of-the-art works when considering human selections as evaluation criterion and (b) comparable results to our method can be achieved only when coverage is not considered as a primary selection aspect.

## References

1. Foundations of forgetting and remembering final report. deliverable 2.4 for eu fp7: Project forgetit - concise preservation by combining managed forgetting and contextualization remembering. https://www.forgetit-project.eu/fileadmin/fm-dam/deliverables/ForgetIT˙WP2˙D2.4.new.pdf. Accessed 10 07 2017
2. Foundations of forgetting and remembering intermediate report. deliverable 2.3 for eu fp:7 project forgetit - concise preservation by combining managed forgetting and contextualization remembering. https://www.forgetit-project.eu/fileadmin/fm-dam/deliverables/ForgetIT˙WP2˙D2.3.pdf. Accessed 10 07 2017
3. Afrasiabi Rad, P., Nilsson, J., Päivärinta, T.: Administration of Digital Preservation Services in the Cloud Over Time: Design Issues and Challenges for Organizations. In: B. Endicott-Popovsky (ed.) The Proceedings of the 2nd International Conference on Cloud Security Management, Proceedings of the international conference on cloud security management. Academic Conferences and Publishing International Ltd (2014)
4. Agrawal, R., Gollapudi, S., Halverson, A., Ieong, S.: Diversifying search results. In: Proceedings of the Second ACM International Conference on Web Search and Data Mining, WSDM '09, pp. 5–14. ACM, New York, NY, USA (2009). DOI 10.1145/1498759.1498766. URL http://doi.acm.org/10.1145/1498759.1498766
5. Szegedy et al., C.: Going deeper with convolutions. In: CVPR 2015 (2015). URL http://arxiv.org/abs/1409.4842
6. Markatopoulou et al., F.: A study on the use of a binary local descriptor and color extensions of local descriptors for video concept detection. In: MultiMedia Modeling Conf., *LNCS*, vol. 8935, pp. 282–293. Springer (2015)
7. Over et al., P.: Trecvid 2013 – an overview of the goals, tasks, data, evaluation mechanisms and metrics. In: Proc. of TRECVID 2013. NIST, USA (2013)
8. Alberdi, E.: Accommodating surprise in taxonomic tasks: a psychological and computational investigation. Ph.D. thesis, University of Aberdeen (1996)
9. Alberdi, E., Sleeman, D.H.: Retax: a step in the automation of taxonomic revision. Artificial Intelligence **91**(2), 257–279 (1997)
10. Alexe, B., Deselaers, T., Ferrari, V.: Measuring the objectness of image windows. Pattern Analysis and Machine Intelligence, IEEE Transactions on **34**(11), 2189–2202 (2012)
11. Ames, M., Eckles, D., Naaman, M., Spasojevic, M., Van House, N.: Requirements for mobile photoware. Personal and Ubiquitous Computing **14**(2), 95–109 (2010)
12. Anastasiu, D.C., Karypis, G.: L2ap: Fast cosine similarity search with prefix l-2 norm bounds. In: 30th Int. Conf on Data Engineering (ICDE), pp. 784–795. IEEE (2014)
13. Anderson, J.R., Schooler, L.J.: Reflections of the environment in memory. Psychological Science **6**(2), 396–408 (1991)
14. Anderson, M.C., Baddeley, A., Eysenck, M., Anderson, M.: Retrieval. Memory pp. 163–189 (2009)
15. Andersson, I., Lindqvist, G., Randers, F.: Preservation Services Planning: a Decision Support Framework. In: The Proceedings of the 2nd International Conference on Cloud Security Management (2014)
16. Andoni, A.: Locality-sensitive hashing scheme based on p-stable distributions. Nearest neighbor methods in learning and vision: Theory and practice (2006)
17. Apostolidis, E., Mezaris, V.: Fast shot segmentation combining global and local visual descriptors. In: Proc. of Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), pp. 6583–6587. IEEE (2014)
18. Apostolidis, K., Papagiannopoulou, C., Mezaris, V.: CERTH at MediaEval 2014 synchronization of multi-user event media task. In: Proceedings of the MediaEval Workshop (2014)
19. Apostolidis, K., Solachidis, V., Papadopoulou, O., Mezaris, V.: Photo collection contextualization. In: 2015 IEEE International Conference on Multimedia Expo Workshops (ICMEW), pp. 1–6 (2015). DOI 10.1109/ICMEW.2015.7169761

20. Arandjelovic, R., Zisserman, A.: All about vlad. In: Proc. of the Conf. on Computer Vision and Pattern Recognition, pp. 1578–1585. IEEE (2013)
21. Argyriou, A., Evgeniou, T., Pontil, M.: Multi-task feature learning. In: Advances in Neural Information Processing Systems (NIPS 2007). MIT Press (2007)
22. Argyriou, A., Evgeniou, T., Pontil, M.: Convex multi-task feature learning. Machine Learning **73**(3), 243–272 (2008)
23. Awad, M., Khan, L., Thuraisingham, B.: Predicting www surfing using multiple evidence combination. The VLDB Journal **17**(3), 401–417 (2008). DOI 10.1007/s00778-006-0014-1. URL http://dx.doi.org/10.1007/s00778-006-0014-1
24. Babenko, A., Lempitsky, V.: Aggregating local deep features for image retrieval. In: Proc. of the Int. Conf. on Computer Vision (ICCV), pp. 1269–1277. IEEE (2015)
25. Backstrom, L., Leskovec, J.: Supervised random walks: Predicting and recommending links in social networks. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, WSDM '11, pp. 635–644. ACM, New York, NY, USA (2011). DOI 10.1145/1935826.1935914. URL http://doi.acm.org/10.1145/1935826.1935914
26. Baddeley, A.: Essentials of human memory. Psychology Press pp. 139–144 (1999)
27. Baddeley, A., Eysenck, M., Anderson, M.: Memory. Psychology Press (2009)
28. Baddeley, A., Eysenck, M., Anderson, M.: Memory. Taylor & Francis (2015). URL https://books.google.gr/books?id=nxNWBQAAQBAJ
29. Baddeley, A., Logie, R.: Working memory: The multiple-component model. in, a. miyake, & p. shah (eds.), models of working memory: Mechanisms of active maintenance and executive control (pp. 28-61) (1999)
30. Baddeley, A.D., Hitch, G.: Working memory. Psychology of learning and motivation **8**, 47–89 (1974)
31. Baeza-Yates, R., Gionis, A., Junqueira, F., Murdock, V., Plachouras, V., Silvestri, F.: The impact of caching on search engines. In: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '07, pp. 183–190. ACM, New York, NY, USA (2007). DOI 10.1145/1277741.1277775. URL http://doi.acm.org/10.1145/1277741.1277775
32. Bahrick, H.P.: Semantic memory content in permastore: Fifty years of memory for spanish learned in school. Journal of experimental psychology: General **113**(1), 1 (1984)
33. Banks, R., Kirk, D., Sellen, A.: A design perspective on three technology heirlooms. Human–Computer Interaction **27**(1-2), 63–91 (2012)
34. Barthel, R., Mackley, K.L., Hudson-Smith, A., Karpovich, A., de Jode, M., Speed, C.: An internet of old things as an augmented memory system. Personal and ubiquitous computing **17**(2), 321–333 (2013)
35. Bartlett, F.C., Burt, C.: Remembering: A study in experimental and social psychology. British Journal of Educational Psychology **3**(2), 187–192 (1933)
36. Barua, D., Kay, J., Kummerfeld, B., Paris, C.: Theoretical foundations for user-controlled forgetting in scrutable long term user models. In: Proceedings of the 23rd Australian Computer-Human Interaction Conference, pp. 40–49. ACM (2011)
37. Basave, A.E.C., Rizzo, G., Varga, A., Rowe, M., Stankovic, M., Dadzie, A.: Making sense of microposts (#microposts2014) named entity extraction & linking challenge. In: 4th Workshop on Making Sense of Microposts (2014)
38. Bendersky, M., Croft, W.B.: Discovering key concepts in verbose queries. In: Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '08, pp. 491–498 (2008)
39. Berberich, K., Bedathur, S., Alonso, O., Weikum, G.: A language modeling approach for temporal information needs. In: Proceedings of the 32Nd European Conference on Advances in Information Retrieval, ECIR'2010, pp. 13–25 (2010)
40. Berdan, R.: Composition and the elements of visual design. Photo composition articles (2004)
41. Bernardi, A., Grimnes, G.A., Groza, T., Scerri, S.: The NEPOMUK Semantic Desktop. In: P. Warren, J. Davies, E. Simperl (eds.) Context and Semantics for Knowledge Management - Technologies for Personal Productivity, chap. The NEPOMUK Semantic Desktop, pp. 255–274. Springer (2011)

42. Berntsen, D., Rubin, D.C.: Cultural life scripts structure recall from autobiographical memory. Memory & cognition **32**(3), 427–442 (2004)
43. Bhattacharya, S., Nojavanasghari, B., Chen, T., Liu, D., Chang, S.F., Shah, M.: Towards a comprehensive computational model foraesthetic assessment of videos. In: Proc. of the 21st Int. Conf. on Multimedia, pp. 361–364. ACM (2013)
44. Blanken, H.M., de Vries, A.P., Blok, H.E., Feng, L.: Multimedia Retrieval. Springer Berlin Heidelberg, NY (2005)
45. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. J. Mach. Learn. Res. (2003)
46. Bontcheva, K., Derczynski, L., Funk, A., Greenwood, M.A., Maynard, D., Aswani, N.: TwitIE: An open-source information extraction pipeline for microblog text. In: Proceedings of the International Conference on Recent Advances in Natural Language Processing. Association for Computational Linguistics (2013)
47. Bordes, A., Ertekin, S., Weston, J., Bottou, L.: Fast kernel classifiers with online and active learning. Journal of Machine Learning Research **6**(Sep), 1579–1619 (2005)
48. Borghoff, U., Rdig, P., Scheffczyk, J.: Long-Term Preservation of Digital Documents, Principles and Practices, online-ausg. edn. Springer-Verlag Berlin Heidelberg (2006). URL http://dx.doi.org/10.1007/978-3-540-33640-2. By Uwe M. Borghoff, Peter Rdig, Lothar Schmitz, Jan Scheffczyk
49. Bos, E.: We can make forgetting impossible, but should we? interactions **2**(3), 11–14 (1995)
50. Bote, J., Fernandez-Feijoo, B., Ruiz, S.: The Cost of Digital Preservation: a Methodological Analysis. Procedia Technology **5**, 103–111 (2012)
51. Bottou, L.: Large-scale machine learning with stochastic gradient descent. In: Proceedings of COMPSTAT, pp. 177–186. Springer (2010)
52. Brand, S.: Escaping the digital dark age. Library Journal **124**(2), 4669 (1999)
53. Bransford, J.: Human cognition: Learning, understanding, and remembering. Thomson Brooks/Cole (1979)
54. Brewer, W.F.: Qualitative analysis of the recalls of randomly sampled autobiographical events. (1988)
55. Brighton, H., Mellish, C.: Advances in instance selection for instance-based learning algorithms. Data mining and knowledge discovery **6**(2), 153–172 (2002)
56. Brown, R., Kulik, J.: Flashbulb memories. Cognition **5**(1), 73–99 (1977)
57. Broyden, C.G.: The convergence of a class of double rank minimization algorithms 1. general considerations. Journal of Applied Mathematics **6**(1), 76–90 (1970)
58. Brubaker, J.R., Callison-Burch, V.: Legacy contact: Designing and implementing post-mortem stewardship at facebook. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, CHI '16, pp. 2908–2919. ACM, New York, NY, USA (2016). DOI 10.1145/2858036.2858254. URL http://doi.acm.org/10.1145/2858036.2858254
59. Cabeza, R., St Jacques, P.: Functional neuroimaging of autobiographical memory. Trends in cognitive sciences **11**(5), 219–227 (2007)
60. Camargo, J.E., González, F.A.: Multimodal latent topic analysis for image collection summarization. Information Sciences **328**, 270–287 (2016)
61. Campos, V., Salvador, A., Giro-i Nieto, X., Jou, B.: Diving deep into sentiment: Understanding fine-tuned cnns for visual sentiment prediction. In: Proceedings of the 1st International Workshop on Affect and Sentiment in Multimedia, ASM '15, pp. 57–62. ACM, New York, NY, USA (2015)
62. Carmel, D., Cohen, D., Fagin, R., Farchi, E., Herscovici, M., Maarek, Y.S., Soffer, A.: Static index pruning for information retrieval systems. In: Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '01, pp. 43–50. ACM, New York, NY, USA (2001). DOI 10.1145/383952.383958. URL http://doi.acm.org/10.1145/383952.383958
63. Carmel, D., Kurland, O.: Query performance prediction for ir. In: Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '12, pp. 1196–1197 (2012)

64. Carmel, D., Yom-Tov, E.: Estimating the query difficulty for information retrieval. In: Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '10, pp. 911–911 (2010)
65. Carpenter, G.A., Grossberg, S.: Adaptive resonance theory. Springer (2017)
66. Cauwenberghs, G., Poggio, T.: Incremental and decremental support vector machine learning. In: Proceedings of the 13th International Conference on Neural Information Processing Systems, NIPS'00, pp. 388–394 (2000)
67. CCSDS: Producer-Archive Interface Methodology Abstract Standard (PAIMAS) - Recommended Practice, CCSDS 651.0-M-1 (Magenta Book) Issue 1. Equivalent to ISO 20652:2006. http://public.ccsds.org/publications/archive/651x0m1.pdf (2004). Retrieved September 2017
68. CCSDS: Reference Model for an Open Archival Information System (OAIS) - Recommended Practice, CCSDS 650.0-M-2 (Magenta Book) Issue 2. Equivalent to ISO 14721:2012. http://public.ccsds.org/publications/archive/650x0m2.pdf (2012). Retrieved September 2017
69. CCSDS Secretariat, S.C., Office, N.: Reference Model for an Open Archival Information System (OAIS). Space Operations Mission Directorate (2009)
70. Ceroni, A., Solachidis, V., Fu, M., Kanhabua, N., Papadopoulou, O., Niederée, C., Mezaris, V.: Investigating human behaviors in selecting personal photos to preserve memories. In: Proceedings of the 2015 IEEE International Conference on Multimedia and Expo (ICME) Workshop on Human Memory-Inspired Multimedia Organization and Preservation (HMMP'15) (2015)
71. Ceroni, A., Solachidis, V., Niederée, C., Papadopoulou, O., Kanhabua, N., Mezaris, V.: To keep or not to keep: An expectation–oriented photo selection method for personal photo collections. In: Proceedings of the ACM International Conference on Multimedia Retrieval (ICMR'2015) (2015)
72. Ceroni, A., Tran, N.K., Kanhabua, N., Niederée, C.: Bridging temporal context gaps using time-aware re-contextualization. In: Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR '14, pp. 1127–1130 (2014)
73. Chang, Y.W., Lin, C.J.: Feature ranking using linear svm. In: Proc. of WCCI Causation and Prediction Challenge, pp. 53–64 (2008)
74. Chappell, D.: Enterprise Service Bus. O'Reilly Media, Inc. (2004)
75. Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A.: Return of the devil in the details: Delving deep into convolutional nets. In: British Machine Vision Conference (2014)
76. Chen, J., Hibino, S.: Reminiscing view: Event-based browsing of consumer's photo and video-clip collections. In: Tenth IEEE International Symposium on Multimedia, pp. 23–30 (2008). DOI 10.1109/ISM.2008.104
77. Chen, R.C., Lee, C.J., Croft, W.B.: On divergence measures and static index pruning. In: Proceedings of the 2015 International Conference on The Theory of Information Retrieval, ICTIR '15, pp. 151–160. ACM, New York, NY, USA (2015). DOI 10.1145/2808194.2809472. URL http://doi.acm.org/10.1145/2808194.2809472
78. Chen, Y.: Information valuation for information lifecycle management. In: Proceedings of International Conference on Autonomic Computing (2005)
79. Chu, W.T., Lin, C.H.: Automatic selection of representative photo and smart thumbnailing using near-duplicate detection. In: Proceedings of the 16th ACM International Conference on Multimedia, MM '08, pp. 829–832. ACM (2008)
80. Ciglan, M., Nørvåg, K.: Wikipop: Personalized event detection system based on wikipedia page view statistics. In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management, CIKM '10, pp. 1931–1932 (2010)
81. Cocchini, G., Logie, R.H., Della Sala, S., MacPherson, S.E., Baddeley, A.D.: Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. Memory & Cognition **30**(7), 1086–1095 (2002)

82. Cohen, J., Mihailidis, P.: Storify and news curation: Teaching and learning about digital storytelling. In: Second Annual Social Media Technology Conference & Workshop, vol. 1, pp. 27–31 (2012)
83. Conci, N., De Natale, F.G., Mezaris, V., Matton, M.: Synchronization of multi-user event media (sem) at mediaeval 2014: Task description, datasets, and evaluation. In: MediaEval (2014)
84. Conci, N., Natale, F.D., Mezaris, V.: Synchronization of multi-user event media (sem) at mediaeval 2014: Task description, datasets, and evaluation. In: MediaEval 2014 Workshop, Barcelona, Spain (2014)
85. Library of Congress, M.E.B.: Metadata Encoding & Transmission Standard, Version 1.11. http://www.loc.gov/standards/mets (2015). Retrieved September 2017
86. Library of Congress, P.E.C.: PREMIS Data Dictionary for Preservation Metadata, Version 3.0. http://www.loc.gov/standards/premis/v3/index.html (2015). Retrieved September 2017
87. Connor, M., Spitzer, S.: The places frecency algorithm. https://developer.mozilla.org/en-US/docs/Mozilla/Tech/Places/Frecency˙algorithm (2015 (accessed August 31, 2015))
88. Conway, M.A.: Memory and the self. Journal of memory and language **53**(4), 594–628 (2005)
89. Conway, M.A.: Episodic memories. Neuropsychologia **47**(11), 2305–2313 (2009)
90. Conway, M.A., Anderson, S.J., Larsen, S.F., Donnelly, C.M., McDaniel, M.A., McClelland, A.G., Rawles, R.E., Logie, R.H.: The formation of flashbulb memories. Memory & Cognition **22**(3), 326–343 (1994)
91. Conway, M.A., Cohen, G., Stanhope, N.: Very long-term memory for knowledge acquired at school and university. Applied cognitive psychology **6**(6), 467–482 (1992)
92. Conway, M.A., Loveday, C.: Accessing autobiographical memories. Act of Remembering: Toward an Understanding of How We Recall the Past , The pp. 56–70 (2010)
93. Conway, M.A., Loveday, C.: Accessing Autobiographical Memories, the Act of Remembering: Toward an Understanding of How We Recall the Past. Wiley Blackwell pp. 56–70 (2010)
94. Conway, M.A., Pleydell-Pearce, C.W.: The construction of autobiographical memories in the self-memory system. Psychological review **107**(2), 261 (2000)
95. Conway, P.: Overview: Rationale for Digitization and Preservation, chap. 2, pp. 5–20. NEDCC (2000)
96. Cook, T.: Macroappraisal in theory and practice: Origins, characteristics, and implementation in canada, 19502000. Archival Science **5**(2-4), 101–161 (2005). URL http://dx.doi.org/10.1007/s10502-005-9010-2
97. Cooper, M., Foote, J., Girgensohn, A., Wilcox, L.: Temporal event clustering for digital photo collections. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) **1**(3), 269–288 (2005)
98. Cooper, M., Foote, J., Girgensohn, A., Wilcox, L.: Temporal event clustering for digital photo collections. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) **1**(3), 269–288 (2005)
99. Copeland, A.: The use of personal value estimations to select images for preservation in public library digital community collections. Future Internet **6**(2) (2014)
100. Cosley, D., Sosik, V.S., Schultz, J., Peesapati, S.T., Lee, S.: Experiences with designing tools for everyday reminiscing. Human–Computer Interaction **27**(1-2), 175–198 (2012)
101. da Costa Pinto, A.a.n., Baddeley, A.D.: Where did you park your car? analysis of a naturalistic long-term recency effect. European Journal of Cognitive Psychology **3**(3), 297–313 (1991)
102. Cowan, N.: Working memory capacity. Psychology press (2012)
103. Crestani, F., Lee, P.L.: Searching the web by constrained spreading activation. Information Processing and Management **36**(4), 585 – 605 (2000). DOI http://dx.doi.org/10.1016/S0306-4573(99)00073-4. URL http://www.sciencedirect.com/science/article/pii/S0306457399000734
104. Crovitz, H.F., Schiffman, H.: Frequency of episodic memories as a function of their age. Bulletin of the Psychonomic Society **4**(5), 517–518 (1974)

105. Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: ECCV Workshop on statistical learning in computer vision, vol. 1, pp. 1–2. Prague (2004)
106. Cubelli, R., Della Sala, S.: Flashbulb memories: Special but not iconic. Cortex **44**(7), 908–909 (2008)
107. Cubelli, R., Della Sala, S.: Flashbulb memories: Sorry, no flash! Cortex **1**(49), 356–357 (2013)
108. Cunningham, H., Tablan, V., Roberts, I., Greenwood, M.A., Aswani, N.: Information Extraction and Semantic Annotation for Multi-Paradigm Information Management. In: M. Lupu, K. Mayer, J. Tait, A.J. Trippe (eds.) Current Challenges in Patent Information Retrieval, *The Information Retrieval Series*, vol. 29, pp. 307–327. Springer Berlin Heidelberg (2011). DOI 10.1007/978-3-642-19231-9_15
109. Curci, A., Conway, M.A.: Playing the flashbulb memory game: A comment on cubelli and della sala. Cortex **49**(1), 352–355 (2013)
110. Cutrell, E., Dumais, S.T., Teevan, J.: Searching to eliminate personal information management. Commun. ACM **49**(1), 58–64 (2006)
111. Daelemans, W., Van Den Bosch, A., Zavrel, J.: Forgetting exceptions is harmful in language learning. Machine learning **34**(1-3), 11–41 (1999)
112. Dang, V., Croft, W.B.: Feature selection for document ranking using best first search and coordinate ascent. In: Proc. of SIGIR'10 Workshop on Feature Generation and Selection for Information Retrieval (2010)
113. Das, M., Loui, A.: Detecting significant events in personal image collections. In: Semantic Computing, 2009. ICSC '09. IEEE International Conference on, pp. 116–123 (2009)
114. Datta, R., Li, J., Wang, J.Z.: Studying aesthetics in photographic images using a computational approach (2012). US Patent App. 13/542,326
115. Datta, R., Wang, J.Z.: ACQUINE: aesthetic quality inference engine-real-time automatic rating of photo aesthetics. In: Proceedings of the Int. Conf. on Multimedia information retrieval, pp. 421–424. ACM (2010)
116. Daumé III, H.: Bayesian multitask learning with latent hierarchies. In: Proc. of the 25th Conf. on Uncertainty in Artificial Intelligence (UAI '09), pp. 135–142. AUAI Press, Arlington, Virginia, US (2009)
117. De Groot, A.D.: Thought and choice in chess, vol. 4. Walter de Gruyter GmbH & Co KG (1978)
118. Dengel, A. (ed.): Semantische Technologien. Spektrum (2012)
119. Dhar, S., Ordonez, V., Berg, T.L.: High level describable attributes for predicting aesthetics and interestingness. In: Computer Vision and Pattern Recognition (CVPR),IEEE Conf. on, pp. 1657–1664 (2011)
120. Djafari Naini, K., Altingovde, I.S.: Exploiting result diversification methods for feature selection in learning to rank. In: M. de Rijke, T. Kenter, A.P. de Vries, C. Zhai, F. de Jong, K. Radinsky, K. Hofmann (eds.) Advances in Information Retrieval: 36th European Conference on IR Research, ECIR 2014, Amsterdam, The Netherlands, April 13-16, 2014. Proceedings, pp. 455–461. Springer International Publishing, Cham (2014). DOI 10.1007/978-3-319-06028-6˙41. URL http://dx.doi.org/10.1007/978-3-319-06028-6˙41
121. Djafari Naini, K., Kawase, R., Kanhabua, N., Niederee, C.: Characterizing high-impact features for content retention in social web applications. In: Proceedings of the 23rd International Conference on World Wide Web, WWW '14 Companion, pp. 559–560. ACM, New York, NY, USA (2014). DOI 10.1145/2567948.2576954. URL http://doi.acm.org/10.1145/2567948.2576954
122. Douze, M., Jégou, H., Sandhawalia, H., Amsaleg, L., Schmid, C.: Evaluation of gist descriptors for web-scale image search. In: Proc. of the Int. Conf. on Image and Video Retrieval, p. 19. ACM (2009)
123. Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A.J., Vapnik, V.: Support vector regression machines. In: M.C. Mozer, M.I. Jordan, T. Petsche (eds.) Advances in Neural Information Processing Systems 9, pp. 155–161 (1997)

124. Duan, L.Y., Lin, J., Wang, Z., Huang, T., Gao, W.: Weighted component hashing of binary aggregated descriptors for fast visual search. IEEE Transactions on multimedia **17**(6), 828–842 (2015)
125. Eaton, E., Ruvolo, P.L.: Ella: An efficient lifelong learning algorithm. In: S. Dasgupta, D.Mcallester (eds.) Proc. of the 30th Int. Conf. on Machine Learning (ICML-13), vol. 28, pp. 507–515. JMLR Workshop (2013)
126. Ebbinghaus, H.: Memory: A contribution to experimental psychology. (1964)
127. EBU: Class Conceptual Data Model (CCDM) - TECH 3351. https://tech.ebu.ch/metadata (2016). Retrieved September 2017
128. Eder, J., Koncilia, C.: Modelling Changes in Ontologies. In: Proceedings of the On The Move - Federated Conferences (OTM 2004) (2004)
129. Edworthy, J., Hellier, E.: Fewer but better auditory alarms will improve patient safety. BMJ Quality & Safety **14**(3), 212–215 (2005). DOI 10.1136/qshc.2004.013052. URL http://qualitysafety.bmj.com/content/14/3/212
130. Eiter, T., Wang, K.: Semantic forgetting in answer set programming. Artificial Intelligence **172**(14), 1644–1672 (2008)
131. Eldesouky, B., Bakry, M., Maus, H., Dengel, A.: Supporting Early Contextualization of Textual Content in Digital Documents on the Web. In: Proc. of the 13th Int. Conf. on Document Analysis and Recognition. Nancy, France (2015)
132. Eldesouky, B., Bakry, M., Maus, H., Dengel, A.: Seed, an end-user text composition tool for the semantic web. In: The Semantic Web - ISWC 2016, 15th International Semantic Web Conference, Proceedings. International Semantic Web Conference (ISWC-2016), 15th, October 17-21, Kobe, Japan. Springer (2016)
133. Eldesouky, B., Maus, H., Schwarz, S., Dengel, A.: Seed, a natural language interface to knowledge bases. In: S. Yamamoto (ed.) Human Interface and the Management of Information. Information and Knowledge Design, *LNCS*, vol. 9172, pp. 280–290. Springer (2015)
134. Ericcson, K.A., Chase, W.G., Faloon, S.: Acquisition of a memory skill. Science **208**(4448), 1181–1182 (1980)
135. Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A.: The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results
136. Evgeniou, T., Pontil, M.: Regularized multi–task learning. In: Proc. of the 10th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD ’04), pp. 109–117. ACM, NY, USA (2004)
137. Fagni, T., Perego, R., Silvestri, F., Orlando, S.: Boosting the performance of web search engines: Caching and prefetching query results by exploiting historical usage data. ACM Trans. Inf. Syst. **24**(1), 51–78 (2006). DOI 10.1145/1125857.1125859. URL http://doi.acm.org/10.1145/1125857.1125859
138. Faidi, T., Chaieb, F., Ghorbel, F.: A new multi-resolution affine invariant planar contour descriptor. In: International Conference on Image Analysis and Processing, pp. 494–505. Springer (2015)
139. Fang, Y., Chang, M.W.: Entity linking on microblogs with spatial and temporal signals. Trans. of the Assoc. for Comp. Linguistics **2**, 259–272 (2014)
140. Ferragina, P., Scaiella, U.: Fast and accurate annotation of short texts with wikipedia pages. IEEE Softw. pp. 70–75 (2012)
141. Firth, J.R.: A Synopsis of Linguistic Theory, 1930-1955. Special Volume of The Philological Society (1957)
142. Fischer, P., Dosovitskiy, A., Brox, T.: Descriptor matching with convolutional neural networks: a comparison to sift. arXiv preprint arXiv:1405.5769 (2014)
143. Fitchett, S., Cockburn, A.: Accessrank: Predicting what users will do next. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’12, pp. 2239–2242. ACM, New York, NY, USA (2012). DOI 10.1145/2207676.2208380. URL http://doi.acm.org/10.1145/2207676.2208380
144. Fitzgerald, J.M., Lawrence, R.: Autobiographical memory across the life-span. Journal of Gerontology **39**(6), 692–698 (1984)

145. Forman, G.: An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. **3**, 1289–1305 (2003)
146. Frohlich, D., Fennell, J.: Sound, paper and memorabilia: resources for a simpler digital photography. Personal and Ubiquitous Computing **11**(2), 107–116 (2007)
147. Fu, M., Ceroni, A., Solachidis, V., Niederée, C., Papadopoulou, O., Kanhabua, N., Mezaris, V.: Learning personalized expectation-oriented photo selection models for personal photo collections. In: IEEE International Conference on Multimedia & Expo Workshops, ICME Workshops, pp. 1–6 (2015)
148. Galton, F.: Psychometric experiments. Brain **2**(2), 149–162 (1879)
149. Gargi, U.: Consumer media capture: Time-based analysis and event clustering. HP-Labs Tech Report (2003)
150. Geng, X., Liu, T.Y., Qin, T., Li, H.: Feature selection for ranking. In: Proc. of SIGIR'07, pp. 407–414 (2007)
151. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (2014)
152. Gladney, H.M.: Preserving Digital Information. Springer (2007). DOI 10.1007/978-3-540-37887-7. URL http://dx.doi.org/10.1007/978-3-540-37887-7
153. Glück, J., Bluck, S.: Looking back across the life span: A life story account of the reminiscence bump. Memory & cognition **35**(8), 1928–1939 (2007)
154. Goetz, T.: Harnessing the Power of Feedback Loops. WIRED Magazine (2011). URL http://www.wired.com/magazine/2011/06/ff_feedbackloop/
155. Goldberg, D., Nichols, D., Oki, B.M., Terry, D.: Using collaborative filtering to weave an information tapestry. Commun. ACM **35**(12), 61–70 (1992). DOI 10.1145/138859.138867. URL http://doi.acm.org/10.1145/138859.138867
156. Gorrell, G., Petrak, J., Bontcheva, K.: Using @Twitter conventions to improve #lod-based named entity disambiguation. In: The Semantic Web. Latest Advances and New Domains, pp. 171–186. Springer (2015)
157. Gowans, G., Campbell, J., Alm, N., Dye, R., Astell, A., Ellis, M.: Designing a multimedia conversation aid for reminiscence therapy in dementia care environments. In: CHI'04 Extended Abstracts on Human Factors in Computing Systems, pp. 825–836. ACM (2004)
158. Grand Challenge at ACM Multimedia Conf. (MM'13): NHK Where is beauty? http://acmmm13.org/subm-issions/call-for-multimedia-grand-challenge-solutions/ task-where-is-beauty/ (2013)
159. Greenberg, D.L.: President bush's false [flashbulb] memory of 9/11/01. Applied Cognitive Psychology **18**(3), 363–370 (2004)
160. Grimnes, G.A., Adrian, B., Schwarz, S., Maus, H., Schumacher, K., Sauermann, L.: Semantic desktop for the end-user. i-com **8**(3), 25–32 (2009)
161. Guo, S., Chang, M.W., Kıcıman, E.: To link or not to link? A study on end-to-end tweet entity linking. In: NAACL-HLT 2013, pp. 1020–1030 (2013)
162. Hanley, J.A., McNeil, B.J.: The meaning and use of the area under a receiver operating characteristic (roc) curve. Radiology **143**(1), 29–36 (1982)
163. Harada, S., Naaman, M., Song, Y.J., Wang, Q., Paepcke, A.: Lost in memories: interacting with photo collections on pdas. In: Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries, pp. 325–333. ACM (2004)
164. Harvey, R.: Appraisal and Selection. In: S. Ross, M. Day (eds.) DCC Digital Curation Manual (2006). URL http ://www.dcc.ac.uk/resource/curation-manual/chapters/appraisal-and-selection
165. Harvey, R.: Appraisal and Selection, chap. 1, pp. 6–39. Digital Curation Center (2007)
166. Harvey, R., Thompson, D.: Automating the Appraisal of Digital Materials. Library Hi Tech **28**(2), 313–322 (2010). DOI 10.1108/07378831011047703. URL http://dx.doi.org/10.1108/07378831011047703
167. He, B., Ounis, I.: Inferring query performance using pre-retrieval predictors. In: In Proc. Symposium on String Processing and Information Retrieval (2004)

168. Henkel, L.A.: Point-and-shoot memories: The influence of taking photos on memory for a museum tour. Psychological science **25**(2), 396–402 (2014)
169. Heutelbeck, D., Brunsmann, J., Wilkes, W., Hundsdörfer, A.: Motivations and Challenges for Digital Preservation in Design and Engineering InDP 2009. In: Proc. of the First International Workshop on Innovation in Digital Preservation in conjunction with JCDL 2009. Austin TX, USA (2009)
170. Ho, W.C., Dautenhahn, K., Lim, M.Y., Vargas, P.A., Aylett, R., Enz, S.: An initial memory model for virtual and robot companions supporting migration and long-term interaction. In: Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on, pp. 277–284. IEEE (2009)
171. Ho, W.C., Dautenhahn, K., Nehaniv, C.L.: Computational memory architectures for autobiographic agents interacting in a complex virtual environment: a working model. Connection Science **20**(1), 21–65 (2008)
172. Hodges, S., Berry, E., Wood, K.: Sensecam: A wearable camera that stimulates and rehabilitates autobiographical memory. Memory **19**(7), 685–696 (2011)
173. Hoffart, J., Yosef, M.A., Bordino, I., Fürstenau, H., Pinkal, M., Spaniol, M., Taneva, B., Thater, S., Weikum, G.: Robust disambiguation of named entities in text. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP '11, pp. 782–792 (2011)
174. Hohman, T.J., Peynircioğlu, Z.F., Beason-Held, L.L.: Flexibility of event boundaries in autobiographical memory. Memory **21**(2), 249–260 (2013)
175. Hohpe, G., Woolf, B.: Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA (2003)
176. Hook, E.: Where to Position that Horizon? (Published in Digital Photography School, May, 2013). http://goo.gl/vU1g46
177. van den Hoven, E., Eggen, B.: Informing augmented memory system design through autobiographical memory theory. Personal and Ubiquitous Computing **12**(6), 433–443 (2008)
178. Hutař, J.: Archives New Zealand Migration from Fedora Commons to the Rosetta Digital Preservation System. iPRES2013 Proceedings (2013)
179. Hyun, J., Kim, H.K., Oh, W.G.: Study on performance of mpeg-7 visual descriptors for deformable object retrieval. In: 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV), pp. 1–5. IEEE (2015)
180. Ingwersen, P., Järvelin, K.: The Turn: Integration of Information Seeking and Retrieval in Context (The Information Retrieval Series). Springer-Verlag New York, Inc., Secaucus, NJ, USA (2005)
181. ISO/IEC: FDIS 23000-15 - Information technology – Multimedia application format (MPEG-A) – Part 15: Multimedia preservation application. http://www.iso.org/iso/catalogue˙detail.htm?csnumber=66430 (2016). Retrieved September 2017
182. Isurin, L., Mcdonald, J.L.: Retroactive interference from translation equivalents: Implications for first language forgetting. Memory & cognition **29**(2), 312–319 (2001)
183. Jégou, H., Chum, O.: Negative evidences and co-occurences in image retrieval: The benefit of pca and whitening. Computer Vision–ECCV 2012 pp. 774–787 (2012)
184. Jegou, H., Douze, M., Schmid, C.: Hamming embedding and weak geometric consistency for large scale image search. Computer Vision ECCV pp. 304–317 (2008)
185. Jégou, H., Douze, M., Schmid, C., Pérez, P.: Aggregating local descriptors into a compact image representation. In: Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 3304–3311. IEEE (2010)
186. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM international conference on Multimedia, pp. 675–678. ACM (2014)
187. Jiang, Y.G., Ngo, C.W., Yang, J.: Towards optimal bag-of-features for object categorization and semantic video retrieval. In: Proceedings of the 6th ACM Int. Conf. on Image and Video Retrieval, CIVR '07, pp. 494–501. ACM, NY, USA (2007)

188. Jilek, C., Maus, H., Schwarz, S., Dengel, A.: Diary generation from personal information models to support contextual remembering and reminiscence. In: IEEE Int. Conf. on Multimedia & Expo Workshops, ICME Workshops 2015, Turin, Italy, pp. 1–6 (2015)
189. Jinda-Apiraksa, A., Vonikakis, V., Winkler, S.: California-nd: An annotated dataset for near-duplicate detection in personal photo collections. In: 5th Int. Workshop on Quality of Multimedia Experience (QoMEX), pp. 142–147. IEEE (2013)
190. Joachims, T.: Optimizing search engines using clickthrough data. In: Proc. of KDD'02, pp. 133–142 (2002)
191. John, J.L., Rowlands, I., Williams, P., Dean, K.: Digital lives: Personal digital archives for the 21st century – an initial synthesis. research paper Beta Version 0.2, British Library (2010). http://britishlibrary.typepad.co.uk/files/digital-lives-synthesis02-1.pdf last accessed October 16, 2016
192. John Walker, S.: Big data: A revolution that will transform how we live, work, and think (2014)
193. Jones, W.: Keeping Found Things Found: The Study and Practice of Personal Information Management. Morgan Kaufmann (2007)
194. Julius, M.: Themen in Sozialen Netzwerken (topics in social networks). blog (2013). URL http://multimedia-memories.tumblr.com/post/45185122672/. Last accessed October 16, 2017
195. Kalnikaitė, V., Whittaker, S.: A saunter down memory lane: Digital reflection on personal mementos. International Journal of Human-Computer Studies **69**(5), 298–310 (2011)
196. Kanhabua, N., Niedere, C., Siberski, W.: Towards concise preservation by managed forgetting: Research issues and case study. In: Proceedings of the 10th International Conference on Preservation of Digital Objects (iPres) (2013)
197. Kanhabua, N., Nørvåg, K.: Determining time of queries for re-ranking search results. In: Proceedings of the 14th European Conference on Research and Advanced Technology for Digital Libraries, ECDL'10, pp. 261–272 (2010)
198. Kao, S.H., Day, W.Y., Cheng, P.J.: An aesthetic-based approach to re-ranking web images. In: Information Retrieval Technology, pp. 610–623. Springer (2010)
199. Kärkkäinen, T., Kaakinen, A., Vainio, T., Väänänen-Vainio-Mattila, K.: Design guidelines for managing metadata for personal content on mobile devices. In: Proceedings of the 7th International Conference on Mobile and Ubiquitous Multimedia, pp. 162–167. ACM (2008)
200. Karpicke, J.D., Roediger, H.L.: Repeated retrieval during learning is the key to long-term retention. Journal of Memory and Language **57**(2), 151–162 (2007)
201. Katifori, A., Vassilakis, C., Dix, A.: Ontologies and the brain: Using spreading activation through ontologies to support personal interaction. Cognitive Systems Research **11**(1), 25 – 41 (2010). DOI http://dx.doi.org/10.1016/j.cogsys.2009.02.001. URL http://www.sciencedirect.com/science/article/pii/S1389041709000138. Brain Informatics
202. Kawase, R., Papadakis, G., Herder, E., Nejdl, W.: Beyond the usual suspects: Context-aware revisitation support. In: ACM Hypertext, pp. 27–36 (2011). DOI 10.1145/1995966.1995974. URL http://doi.acm.org/10.1145/1995966.1995974
203. Kaya Kızılöz, B., Tekcan, A.I.: Canonical categories in flashbulb memories. Applied Cognitive Psychology **27**(3), 352–359 (2013)
204. Kaye, J.J., Vertesi, J., Avery, S., Dafoe, A., David, S., Onaga, L., Rosero, I., Pinch, T.: To have and to hold: Exploring the personal archive. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '06, pp. 275–284. ACM, New York, NY, USA (2006). DOI 10.1145/1124772.1124814. URL http://doi.acm.org/10.1145/1124772.1124814
205. Ke, Y., Tang, X., Jing, F.: The design of high-level features for photo quality assessment. In: Computer Vision and Pattern Recognition, IEEE Computer Society Conf. on, vol. 1, pp. 419–426 (2006)
206. Kejser, U.B., Nielsen, A.B., Thirifays, A.: Cost Model for Digital Preservation: Cost of Digital Migration. IJDC **6**(1), 255–267 (2011)
207. Kelly, P., Marshall, S.J., Badland, H., Kerr, J., Oliver, M., Doherty, A.R., Foster, C.: An ethical framework for automated, wearable cameras in health behavior research. American journal of preventive medicine **44**(3), 314–319 (2013)

208. Kesireddi, R.C., Raj, G.: Embedding, extracting and matching of fingerprint images using digital watermarking. In: Proc. of the Int. Conf. on Advanced Research in Computer Science Engineering & Technology (ICARCSET), ICARCSET '15. ACM, New York, NY, USA (2015)
209. Kimeldorf, G., Wahba, G.: Some results on Tchebycheffian spline functions. Journal of Mathematical Analysis and Applications **33**(1), 82–95 (1971)
210. Kindberg, T., Spasojevic, M., Fleck, R., Sellen, A.: How and why people use camera phones. HP Laboratories Technical Report HPL-2004-216 (2004)
211. Kirk, D., Sellen, A., Rother, C., Wood, K.: Understanding photowork. In: Proceedings of the SIGCHI conference on Human Factors in computing systems, pp. 761–770. ACM (2006)
212. Koppel, J., Brown, A.D., Stone, C.B., Coman, A., Hirst, W.: Remembering president barack obama's inauguration and the landing of us airways flight 1549: A comparison of the predictors of autobiographical and event memory. Memory **21**(7), 798–806 (2013)
213. Koutstaal, W., Schacter, D.L., Johnson, M.K., Angell, K.E., Gross, M.S.: Post-event review in older and younger adults: improving memory accessibility of complex everyday events. Psychology and aging **13**(2), 277 (1998)
214. Koutstaal, W., Schacter, D.L., Johnson, M.K., Galluccio, L.: Facilitation and impairment of event memory produced by photograph review. Memory & Cognition **27**(3), 478–493 (1999)
215. Krizhevsky, A., Ilya, S., Hinton, G.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems (NIPS 2012), pp. 1097–1105. Curran Associates, Inc. (2012)
216. Kumar, A., Daume, H.: Learning task grouping and overlap in multi-task learning. In: J. Langford, J. Pineau (eds.) Proc. of the 29th Int. Conf. on Machine Learning (ICML-12), pp. 1383–1390. ACM, NY, USA (2012)
217. Kuny, T.: A digital dark ages? challenges in the preservation of electronic information. In: 63RD IFLA Council and General Conference, Workshop on Audiovisual and Multimedia joint with Preservation and Conservation, InformationTechnology, Library Buildings and Equipment, and the PAC Core Programme (1997)
218. Kuny, T.: The Digital Dark Ages? Challenges in the Preservation of Electronic Information. International Preservation News (17) (1998). URL http://www.ifla.org/VI/4/news/17-98.htm#2
219. Kurby, C.A., Zacks, J.M.: Segmentation in the perception and memory of events. Trends in cognitive sciences **12**(2), 72–79 (2008)
220. Kuwahara, N., Abe, S., Yasuda, K., Kuwabara, K.: Networked reminiscence therapy for individuals with dementia by using photo and video sharing. In: Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility, pp. 125–132. ACM (2006)
221. Kvavilashvili, L., Mirani, J., Schlagman, S., Kornbrot, D.E.: Comparing flashbulb memories of september 11 and the death of princess diana: Effects of time delays and nationality. Applied Cognitive Psychology **17**(9), 1017–1031 (2003)
222. Lakatos, I.: Proofs and refutations (i). The British Journal for the Philosophy of science **14**(53), 1–25 (1963)
223. Lampasona, C., Rostanin, O., Maus, H.: Seamless integration of order processing in ms outlook using smartoffice: an empirical evaluation. In: ESEM '12 Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement. ACM Press (2012). DOI 10.1145/2372251.2372281
224. Landauer, T.K., Bjork, R.A.: Optimum rehearsal patterns and name learning. Practical aspects of memory **1**, 625–632 (1978)
225. Lang, J., Marquis, P.: Reasoning under inconsistency: A forgetting-based approach. Artificial Intelligence **174**(12-13), 799–823 (2010)
226. Lantz, B.: Photography Composition Tips and Techniques (Published in Macro Photography, 2013). http://goo.gl/fsx20y
227. Lappas, T., Arai, B., Platakis, M., Kotsakos, D., Gunopulos, D.: On burstiness-aware search for document sequences. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '09, pp. 477–486 (2009)

228. Lavoie, B., Dempsey, L.: Thirteen ways of looking at...digital preservation. D-Lib Magazine **10**(7/8) (2004). URL http://dblp.uni-trier.de/db/journals/dlib/dlib10.html
229. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)
230. Lee, C.A. (ed.): I, Digital: Personal Collections in the Digital Era. Society of American Archivists (2011)
231. Lee, C.J., Chen, R.C., Kao, S.H., Cheng, P.J.: A term dependency-based approach for query terms ranking. In: Proceedings of the 18th ACM Conference on Information and Knowledge Management, CIKM '09, pp. 1267–1276 (2009)
232. Lehmann, J., Gonçalves, B., Ramasco, J.J., Cattuto, C.: Dynamical classes of collective attention in twitter. In: Proceedings of the 21st International Conference on World Wide Web, WWW '12, pp. 251–260 (2012)
233. Lei, Y., Zheng, L., Huang, J.: Geometric invariant features in the radon transform domain for near-duplicate image detection. Pattern Recognition Letters **47**(11), 3630–3640 (2014)
234. Levine, B., Svoboda, E., Hay, J.F., Winocur, G., Moscovitch, M.: Aging and autobiographical memory: dissociating episodic from semantic retrieval. Psychology and aging **17**(4), 677 (2002)
235. Li, C., Loui, A.C., Chen, T.: Towards aesthetics: a photo quality assessment and photo selection system. In: Proceedings of the Int. Conf. on Multimedia, pp. 827–830. ACM (2010)
236. Li, J., Lim, J.H., Tian, Q.: Automatic summarization for personal digital photos. In: Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint, vol. 3, pp. 1536–1540 vol.3 (2003)
237. Li, X., Larson, M., Hanjalic, A.: Pairwise geometric matching for large-scale object retrieval. In: Proc. of the Conf. on Computer Vision and Pattern Recognition, pp. 5153–5161. IEEE (2015)
238. Li, Y., Sheopuri, A.: Applying image analysis to assess food aesthetics and uniqueness. In: Proc. of Int. Conf. on Image Processing (ICIP), pp. 311–314. IEEE (2015)
239. Library of Congress: Personal archiving: Preserving your digital memories. URL http://www.digitalpreservation.gov/personalarchiving. Last accessed: October 16, 2017
240. Library of Congress: Preserving digital memories. In: 7th International Conference on Preservation of Digital Objects (iPRES 2010) (2010). http://www.digitalpreservation.gov/personalarchiving/documents/NDIIP˙PA˙poster.pdf last accessed: October 16, 2017
241. Lim, M.Y.: Memory models for intelligent social companions. (2012)
242. Lim, M.Y., Aylett, R., Ho, W.C., Enz, S., Vargas, P.: A socially-aware memory for companion agents. In: International Workshop on Intelligent Virtual Agents, pp. 20–26. Springer (2009)
243. Lim, M.Y., Aylett, R., Vargas, P.A., Ho, W.C., Dias, J.: Human-like memory retrieval mechanisms for social companions. In: The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 3, pp. 1117–1118. International Foundation for Autonomous Agents and Multiagent Systems (2011)
244. Lin, C.Y.: Rouge: A package for automatic evaluation of summaries. In: Text summarization branches out: Proceedings of the ACL-04 workshop, vol. 8 (2004)
245. Lindley, S.E.: Before i forget: From personal memory to family history. Human–Computer Interaction **27**(1-2), 13–36 (2012)
246. Lindsay, R.C.: Handbook Of Eyewitness Psychology 2 Volume Set. Psychology Press (2014)
247. Linton, M.: Memory for real-world events. Explorations in cognition pp. 376–404 (1975)
248. Liu, D.C., Nocedal, J.: On the limited memory BFGS method for large scale optimization. Mathematical programming **45**(1-3), 503–528 (1989)
249. Liu, J., Huang, Z., Cai, H., Shen, H.T., Ngo, C.W., Wang, W.: Near-duplicate video retrieval: Current research and future trends. Computing Surveys (CSUR) **45**(4), 44 (2013)
250. Liu, T.Y.: Learning to rank for information retrieval. Foundations and Trends in Information Retrieval **3**(3), 225–331 (2009)
251. Lo, K.Y., Liu, K.H., Chen, C.S.: Assessment of photo aesthetics with efficiency. In: Pattern Recognition (ICPR),21st Int. Conf. on, pp. 2186–2189. IEEE (2012)

252. Lo, K.Y., Liu, K.H., Chen, C.S.: Intelligent photographing interface with on-device aesthetic quality assessment. In: Computer Vision-ACCV 2012 Workshops, pp. 533–544. Springer (2013)
253. Loftus, E.F., Palmer, J.C.: Reconstruction of automobile destruction: An example of the interaction between language and memory. Journal of verbal learning and verbal behavior **13**(5), 585–589 (1974)
254. Logie, R., Hunter, J., McIntosh, N., Gilhooly, K., Alberdi, E., Reiss, J.: Medical cognition and computer support in the intensive care unit: A cognitive engineering approach. Engineering Psychology and Cognitive Ergonomics: Integration of Theory and Application pp. 167–174 (1997)
255. Logie, R., Wright, R., Decker, S.: Recognition memory performance and residential burglary. Applied Cognitive Psychology **6**(2), 109–123 (1992)
256. Logie, R.H.: Spatial and visual working memory: A mental workspace. Psychology of learning and motivation **42**, 37–78 (2003)
257. Logie, R.H.: The functional organization and capacity limits of working memory. Current Directions in Psychological Science **20**(4), 240–245 (2011). DOI 10.1177/0963721411415340. URL http://dx.doi.org/10.1177/0963721411415340
258. Logie, R.H.: Visuo-spatial working memory. Psychology Press (2014)
259. Loui, A., Savakis, A.: Automated event clustering and quality screening of consumer pictures for digital albuming. IEEE Trans. on Multimedia **5**(3), 390–402 (2003). DOI 10.1109/TMM.2003.814723
260. Loui, A.C., Savakis, A.: Automated event clustering and quality screening of consumer pictures for digital albuming. IEEE Transactions on Multimedia **5**(3), 390–402 (2003)
261. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International journal of computer vision **60**(2), 91–110 (2004)
262. Lu, Y., Wan, Y.: PHA: A fast potential-based hierarchical agglomerative clustering method. Pattern Recognition Letters **46**(5), 1227–1239 (2013)
263. Lucero, A., Boberg, M., Uusitalo, S.: Image space: Capturing, sharing and contextualizing personal pictures in a simple and playful way. In: Proceedings of the International Conference on Advances in Computer Enterntainment Technology, ACE '09, pp. 215–222. ACM, New York, NY, USA (2009). DOI 10.1145/1690388.1690424. URL http://doi.acm.org/10.1145/1690388.1690424
264. Luminet, O., Curci, A.: Flashbulb memories: New issues and new perspectives. Psychology Press (2008)
265. Luo, Y., Tang, X.: Photo and video quality evaluation: Focusing on the subject. In: Computer Vision–ECCV, pp. 386–399. Springer (2008)
266. Maguire, E.A.: Neuroimaging studies of autobiographical event memory. Philosophical Transactions of the Royal Society of London B: Biological Sciences **356**(1413), 1441–1451 (2001)
267. Maji, S., Berg, A., Malik, J.: Classification using intersection kernel support vector machines is efficient. In: IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008., pp. 1–8 (2008)
268. Manning, C.D., Raghavan, P., Schütze, H.: Introduction to Information Retrieval. Cambridge University Press, New York, NY, USA (2008)
269. Manning, C.D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S.J., McClosky, D.: The Stanford CoreNLP natural language processing toolkit. In: ACL '14 (2014)
270. Marchesotti, L., Perronnin, F., Larlus, D., Csurka, G.: Assessing the aesthetic quality of photographs using generic image descriptors. In: Computer Vision (ICCV), IEEE Int. Conf. on, pp. 1784–1791 (2011)
271. Marée, R., Geurts, P., Wehenkel, L.: Content-based image retrieval by indexing random subwindows with randomized trees. In: Asian Conference on Computer Vision, pp. 611–620. Springer (2007)
272. Markatopoulou, F., Mezaris, V., Patras, I.: Cascade of classifiers based on binary, non-binary and deep convolutional network descriptors for video concept detection. In: Proc. of the IEEE Int. Conf. on Image Processing (ICIP 2015), pp. 1786–1790 (2015)

273. Markatopoulou, F., Mezaris, V., Patras, I.: Online multi-task learning for semantic concept detection in video. In: IEEE International Conference on Image Processing (ICIP), pp. 186–190 (2016)
274. Markatopoulou, F., Mezaris, V., Pittaras, N., Patras, I.: Local features and a two-layer stacking architecture for semantic concept detection in video. IEEE Transactions on Emerging Topics in Computing **3**(2), 193–204 (2015)
275. Markatopoulou, F., Pittaras, N., Papadopoulou, O., Mezaris, V., Patras, I.: A Study on the Use of a Binary Local Descriptor and Color Extensions of Local Descriptors for Video Concept Detection, pp. 282–293. Springer International Publishing, Cham (2015)
276. Marshall, C.C.: Challenges and Opportunities for Personal Digital Archiving. Digital: Personal Collections in the Digital Era pp. 90–114 (2011)
277. Marshall, C.C.: I, digital: Personal collections in the digital era. Society of American Archivists pp. 90–114 (2011)
278. Mather, M., Carstensen, L.L.: Aging and motivated cognition: The positivity effect in attention and memory. Trends in cognitive sciences **9**(10), 496–502 (2005)
279. Maus, H., Schwarz, S., Dengel, A.: Weaving personal knowledge spaces into office applications. In: Integration of Practice-Oriented Knowledge Technology: Trends and Prospectives, pp. 71–82. Springer (2013)
280. Maus, H., Schwarz, S., Haas, J., Dengel, A.: CONTASK: Context-Sensitive Task Assistance in the Semantic Desktop. In: J. Filipe, J. Cordeiro (eds.) Enterprise Information Systems. 12th International Conference, ICEIS 2010, Funchal-Madeira, Portugal, June 8-12, 2010, Revised Selected Papers, *Lecture Notes in Business Information Processing*, vol. 73, pp. 177–192. Springer (2011). DOI 10.1007/978-3-642-19802-1˙13
281. Mavridaki, E., Mezaris, V.: No-reference blur assessment in natural images using fourier transform and spatial pyramids. In: Proc. of Int. Conf. on Image Processing (ICIP). IEEE (2014)
282. Mavridaki, E., Mezaris, V.: A comprehensive aesthetic quality assessment method for natural images using basic rules of photography. In: Image Processing (ICIP), 2015 IEEE International Conference on, pp. 887–891. IEEE (2015)
283. Maxwell, K.T., Croft, W.B.: Compact query term selection using topically related text. In: Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '13, pp. 583–592 (2013)
284. Mayer-Schönberger, V.: Delete - The Virtue of Forgetting in the Digital Age. Morgan Kaufmann Publishers (2009)
285. Mayer-Schönberger, V.: Delete: The virtue of forgetting in the digital age. Princeton University Press (2011)
286. Maynard, D., Greenwood, M.A.: Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis. In: Proceedings of LREC 2014. Reykjavik, Iceland (2014)
287. Maynard, D., Roberts, I., Greenwood, M.A., Rout, D., Bontcheva, K.: A Framework for Real-time Semantic Social Media Analysis. Web Semantics: Science, Services and Agents on the World Wide Web (2017). DOI https://doi.org/10.1016/j.websem.2017.05.002. URL http://www.sciencedirect.com/science/article/pii/S1570826817300240
288. McCamy, C.S.: Correlated color temperature as an explicit function of chromaticity coordinates. Color Research & Application **17**(2), 142–144 (1992)
289. McDaniel, M.A., Einstein, G.O.: Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. Applied cognitive psychology **14**(7) (2000)
290. McDaniel, M.A., Einstein, G.O.: Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. Applied Cognitive Psychology **14**, 127–144 (2000)
291. McDaniel, M.A., Einstein, G.O.: Prospective memory: An overview and synthesis of an emerging field. Sage Publications (2007)
292. McDaniel, M.A., Einstein, G.O.: Prospective Memory: An Overview and Synthesis of an Emerging Field. SAGE Publications, Inc. (2007)
293. McGee-Lennon, M.R., Wolters, M.K., Brewster, S.: User-centred multimodal reminders for assistive living. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2105–2114. ACM (2011)

294. McGeoch, J.A., McDonald, W.T.: Meaningful relation and retroactive inhibition. The American Journal of Psychology **43**(4), 579–588 (1931)
295. McKENNA, S.P., Glendon, A.: Occupational first aid training: Decay in cardiopulmonary resuscitation (cpr) skills. Journal of Occupational and Organizational Psychology **58**(2), 109–117 (1985)
296. McLaren, K.: Xiiithe development of the cie 1976 (l* a* b*) uniform colour space and colour-difference formula. Journal of the Society of Dyers and Colourists **92**(9), 338–341 (1976)
297. Meeker, M., Wu, L.: Internet trends. Tech. rep., Kleiner Perkins Caufield and Byers (2013)
298. Mei, T., Wang, B., Hua, X.S., Zhou, H.Q., Li, S.: Probabilistic multimodality fusion for event based home photo clustering. In: Multimedia and Expo, 2006 IEEE International Conference on, pp. 1757–1760. IEEE (2006)
299. Meij, E., Weerkamp, W., de Rijke, M.: Adding semantics to microblog posts. In: Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, pp. 563–572 (2012)
300. Meyer, J.: Emotional Images: how to add feeling with muted tones and harmonious color (Published in Digital Camera World, March 28, 2013). http://goo.gl/7uV9iz
301. Mezaris, V., Sidiropoulos, P., Dimou, A., Kompatsiaris, I.: On the use of visual soft semantics for video temporal decomposition to scenes. In: Semantic Computing (ICSC), 2010 IEEE Fourth International Conference on, pp. 141–148 (2010). DOI 10.1109/ICSC.2010.23
302. Mihalcea, R., Csomai, A.: Wikify!: Linking documents to encyclopedic knowledge. In: Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management, CIKM '07, pp. 233–242 (2007)
303. Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Van Gool, L.: A comparison of affine region detectors. Int. Journal of Computer Vision **65**(1-2), 43–72 (2005)
304. Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological review **63**(2), 81 (1956)
305. Milne, D., Witten, I.H.: Learning to link with wikipedia. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM '08, pp. 509–518 (2008)
306. Mirzasoleiman, B., Badanidiyuru, A., Karbasi, A.: Fast constrained submodular maximization: Personalized data summarization. In: Proceedings of the 33nd International Conference on Machine Learning, ICML, pp. 1358–1367 (2016)
307. Mitra, S., Winslett, M., Hsu, W.W.: Query-based partitioning of documents and indexes for information lifecycle management. In: Proceedings of the 2008 ACM SIGMOD international conference on Management of data, SIGMOD '08, pp. 623–636 (2008). DOI 10.1145/1376616.1376680. URL http://doi.acm.org/10.1145/1376616.1376680
308. Mols, I., Hoven, E.v.d., Eggen, B.: Making memories: A cultural probe study into the remembering of everyday life. In: Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational, NordiCHI '14, pp. 256–265. ACM (2014)
309. Moorthy, A.K., Obrador, P., Oliver, N.: Towards computational models of the visual aesthetic appeal of consumer videos. In: Proc. of the 11th European Conference on Computer Vision (ECCV), Heraklion, Crete, Greece, 2010, pp. 1–14. Springer (2010)
310. Morris, P.E., Tweedy, M., Gruneberg, M.M.: Interest, knowledge and the memorizing of soccer scores. British Journal of Psychology **76**(4), 415–425 (1985)
311. Mothe, J., Tanguy, L.: Linguistic features to predict query difficulty. In: In ACM SIGIR 2005 Workshop on Predicting Query Difficulty - Methods and Applications (2005)
312. Mousavi, H., Srinivas, U., Monga, V., Suo, Y., Dao, M., Tran, T.: Multi-task image classification via collaborative, hierarchical spike-and-slab priors. In: Proc. of the IEEE Int. Conf. on Image Processing (ICIP 2014), pp. 4236–4240 (2014)
313. Murray, N., Marchesotti, L., Perronnin, F.: AVA: A large-scale database for aesthetic visual analysis. In: Computer Vision and Pattern Recognition (CVPR), IEEE Conf. on, pp. 2408–2415 (2012)
314. Murray, N., Marchesotti, L., Perronnin, F., Meylan, F.: Learning to rank images using semantic and aesthetic labels. In: BMVC, pp. 1–10 (2012)

315. Nathan, L., Shaffer, E.: Preserving social media: Opening a multi-disciplinary dialogue. In: Proceedings of the The Memory of the World in the Digital Age: Digitization and Preservation. Vancouver, BC, 26-28 September 2012. UNESCO (2013)
316. Ng, J., Yang, F., Davis, L.: Exploiting local features from deep networks for image retrieval. In: Proc. of the Conf. on Computer Vision and Pattern Recognition Workshops, pp. 53–61. IEEE (2015)
317. Niederee, C., Kanhabua, N., Gallo, F., Logie, R.H.: Forgetful Digital Memory: Towards Brain-Inspired Long-Term Data and Information Management. SIGMOD Rec. **44**(2), 41–46 (2015). DOI 10.1145/2814710.2814718. URL http://doi.acm.org/10.1145/2814710.2814718
318. Niu, Y., Liu, F.: What makes a professional video? a computational aesthetics approach. IEEE Transactions on Circuits and Systems for Video Technology **22**(7), 1037–1049 (2012)
319. OASIS: Content Management Interoperability Services (CMIS) Version 1.1. OASIS Standard. http://docs.oasis-open.org/cmis/CMIS/v1.1/CMIS-v1.1.html (2013). Retrieved September 2017
320. Obozinski, G., Taskar, B.: Multi-task feature selection. In: Proc. of the 23rd Int. Conf. on Machine Learning (ICML 2006). Workshop of Structural Knowledge Transfer for Machine Learning. Pittsburgh, Pennsylvania (2006)
321. Obrador, P., De Oliveira, R., Oliver, N.: Supporting personal photo storytelling for social albums. In: Proceedings of the 18th ACM international conference on Multimedia, pp. 561–570. ACM (2010)
322. Oliva, A., Torralba, A.: Modeling the shape of the scene: A holistic representation of the spatial envelope. Int. journal of computer vision **42**(3), 145–175 (2001)
323. Oquab, M., Bottou, L., Laptev, I., Sivic, J.: Learning and transferring mid-level image representations using convolutional neural networks. In: CVPR (2014)
324. Owoputi, O., O'Connor, B., Dyer, C., Gimpel, K., Schneider, N., Smith, N.A.: Improved part-of-speech tagging for online conversational text with word clusters. In: NAACL-HLT 2013, pp. 380–390 (2013)
325. Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: Bringing order to the web. In: Proceedings of the 7th International World Wide Web Conference, pp. 161–172. Brisbane, Australia (1998). URL citeseer.nj.nec.com/page98pagerank.html
326. Päivärinta, T., Westerlund, P., Nilsson, J.: Integrating Contemporary Content Management and Long-term Digital Preservation: a Design Problem. Springer Verlag, Luleå University of Technology, Information Systems (2015)
327. Papadakis, G., Kawase, R., Herder, E., Niederée, C.: A layered approach to revisitation prediction. In: Proceedings of the 11th International Conference on Web Engineering, ICWE'11, pp. 258–273. Springer-Verlag, Berlin, Heidelberg (2011). URL http://dl.acm.org/citation.cfm?id=2027776.2027796
328. Papagiannopoulou, C., Mezaris, V.: Concept-based image clustering and summarization of event-related image collections. In: Proceedings of the 1st ACM International Workshop on Human Centered Event Understanding from Multimedia, pp. 23–28. ACM (2014)
329. Peetz, M.H., de Rijke, M.: Cognitive temporal document priors. In: Proceedings of the 35th European Conference on Advances in Information Retrieval, ECIR'13, pp. 318–330. Springer-Verlag, Berlin, Heidelberg (2013). DOI 10.1007/978-3-642-36973-5˙27. URL http://dx.doi.org/10.1007/978-3-642-36973-5˙27
330. Perronnin, F., Sánchez, J., Mensink, T.: Improving the fisher kernel for large-scale image classification. Computer Vision ECCV pp. 143–156 (2010)
331. Petrelli, D., van den Hoven, E., Whittaker, S.: Making history: Intentional capture of future memories. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '09, pp. 1723–1732. ACM, New York, NY, USA (2009). DOI 10.1145/1518701.1518966. URL http://doi.acm.org/10.1145/1518701.1518966
332. Petrelli, D., Whittaker, S.: Family memories in the home: contrasting physical and digital mementos. Personal and Ubiquitous Computing **14**(2), 153–169 (2010)
333. Pezdek, K.: Event memory and autobiographical memory for the events of september 11, 2001. Applied Cognitive Psychology **17**(9), 1033–1045 (2003)

334. Philbin, J., Chum, O., Isard, M., Sivic, J., Zisserman, A.: Object retrieval with large vocabularies and fast spatial matching. In: Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 1–8. IEEE (2007)
335. Piolino, P., Desgranges, B., Eustache, F.: Episodic autobiographical memories over the course of time: cognitive, neuropsychological and neuroimaging findings. Neuropsychologia **47**(11), 2314–2329 (2009)
336. Pittaras, N., Markatopoulou, F., Mezaris, V., et al.: Comparison of Fine-Tuning and Extension Strategies for Deep Convolutional Neural Networks, pp. 102–114. Springer, Cham (2017)
337. Ponte, J.M., Croft, W.B.: A language modeling approach to information retrieval. In: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '98, pp. 275–281 (1998)
338. POWRR: Tool-grid (2013). URL http://digitalpowrr.niu.edu/tool-grid/. http://digitalpowrr.niu.edu/tool-grid/, last accessed: October 16, 2017
339. Pritchett, A.R.: Aviation automation: General perspectives and specific guidance for the design of modes and alerts. Reviews of Human Factors and Ergonomics **5**(1), 82–113 (2009). DOI 10.1518/155723409X448026. URL https://doi.org/10.1518/155723409X448026
340. Rabbath, M., Sandhaus, P., Boll, S.: Automatic creation of photo books from stories in social media. In: Proceedings of Second ACM SIGMM Workshop on Social Media, WSM '10, pp. 15–20. ACM (2010)
341. Raiber, F., Kurland, O.: Query-performance prediction: Setting the expectations straight. In: Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR '14, pp. 13–22 (2014)
342. Rauch, J., Braud, M., Edelstein, O., Rabinovici-Cohen, S., Nagin, K., Marberg, J., Voets, D., Sanya, I., Badawy, M., Shebab, E., Randers, F., Droppert, J., Klecha, M.: ENSURE: Long Term Digital Preservation of Health Care, Clinical Trial and Financial data. In: J. Borbinha, M. Nelson, S. Knight (eds.) iPRES 2013: Proceedings of the 10th International Conference on Preservation of Digital Objects (2013)
343. Reason, J.: Human error. Cambridge university press (1990)
344. Renoult, L., Davidson, P.S., Palombo, D.J., Moscovitch, M., Levine, B.: Personal semantics: at the crossroads of semantic and episodic memory. Trends in cognitive sciences **16**(11), 550–558 (2012)
345. Richards, D., Bransky, K.: Forgetmenot: What and how users expect intelligent virtual agents to recall and forget personal conversational content. International Journal of Human-Computer Studies **72**(5), 460–476 (2014)
346. Ritter, A., Clark, S., Mausam, Etzioni, O.: Named entity recognition in tweets: An experimental study. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 1524–1534 (2011)
347. Robertson, S.: A new interpretation of average precision. In: Proc. of the 31st Annual Int. ACM SIGIR conf. on Research and Development in Information Retrieval, pp. 689–690. ACM (2008)
348. Robusto, C.C.: The cosine-haversine formula. The American Mathematical Monthly **64**(1), 38–40 (1957)
349. Rocha, C., Schwabe, D., Aragao, M.P.: A hybrid approach for searching in the semantic web. In: Proceedings of the 13th International Conference on World Wide Web, WWW '04, pp. 374–383. ACM, New York, NY, USA (2004). DOI 10.1145/988672.988723. URL http://doi.acm.org/10.1145/988672.988723
350. Roediger, H.L., Marsh, E.J.: Episodic and autobiographical memory. Handbook of psychology (2003)
351. Russ, A.L., Zillich, A.J., McManus, M.S., Doebbeling, B.N., Saleem, J.J.: Prescribers interactions with medication alerts at the point of prescribing: A multi-method, in situ investigation of the human–computer interaction. International journal of medical informatics **81**(4), 232–243 (2012)
352. Russakovsky, O., Deng, J., et al., H.S.: ImageNet Large Scale Visual Recognition Challenge. Int. Journal of Computer Vision (IJCV) **115**(3), 211–252 (2015). DOI 10.1007/s11263-015-0816-y

353. Samani, Z.R., Moghaddam, M.E.: A knowledge-based semantic approach for image collection summarization. Multimedia Tools and Applications pp. 1–23 (2016)
354. Sansone, E., Apostoldis, K., Conci, N., Boato, G., Mezaris, V., Natale, F.G.B.D.: Automatic synchronization of multi-user photo galleries. IEEE Transactions on Multimedia (2017). DOI 10.1109/TMM.2017.2655446
355. Sauermann, L.: The Gnowsis Semantic Desktop approach to Personal Information Management. Ph.D. thesis, Fachbereich Informatik der Universitt Kaiserslautern (2009). URL http://www.dfki.uni-kl.de/ sauermann/papers/Sauermann2009phd.pdf
356. Sauermann, L., Bernardi, A., Dengel, A.: Overview and Outlook on the Semantic Desktop. In: S. Decker, J. Park, D. Quan, L. Sauermann (eds.) Proceedings of the First Semantic Desktop Workshop at the ISWC Conference 2005 (2005). URL http://www.dfki.uni-kl.de/ sauermann/papers/Sauermann+2005d.pdf
357. Sauermann, L., Dengel, A., van Elst, L., Lauer, A., Maus, H., Schwarz, S.: Personalization in the EPOS project. In: Proceedings of the Semantic Web Personalization Workshop at the ESWC 2006 (2006)
358. Sauermann, L., van Elst, L., Dengel, A.: PIMO – A Framework for Representing Personal Information Models. In: T. Pellegrini, S. Schaffert (eds.) I-SEMANTICS Conference 5-7 September 2007, Graz, Austria, J.UCS, pp. 270–277. Know-Center, Austria (2007)
359. Sauermann, L., Heim, D.: Evaluating long-term use of the gnowsis semantic desktop for pim. In: The Semantic Web. 7th International Semantic Web Conference (ISWC-2008), 7th International Semantic Web Conference, October 26-30, Karlsruhe, Germany, *LNCS*, vol. 5318, pp. 467–482. Springer (2008). URL http://www.springerlink.com/content/c7u686mq27751373/
360. Savakis, A.E., Etz, S.P., Loui, A.C.: Evaluation of image appeal in consumer photography. In: Electronic Imaging, pp. 111–120. International Society for Optics and Photonics (2000)
361. Schacter, D.L., Koutstaal, W., Johnson, M.K., Gross, M.S., Angell, K.E.: False recollection induced by photographs: a comparison of older and younger adults. Psychology and aging **12**(2), 203 (1997)
362. Schellenberg, T.R.: The appraisal of modern records. Bulletins of the National Archives **8**, 46 pages (1956)
363. Schölkopf, B., Herbrich, R., Smola, A.J.: A generalized representer theorem. In: Proc. of 14th Annual Conf. on Computational Learning Theory, pp. 416–426 (2001)
364. Schwarz, S.: Context-Awareness and Context-Sensitive Interfaces for Knowledge Work Support. Phd thesis, University of Kaiserslautern, Fachbereich Informatik (2010). URL http://www.dr.hut-verlag.de/978-3-86853-388-0.html
365. Seah, B.S., Bhowmick, S.S., Sun, A.: Prism: Concept-preserving social image search results summarization. In: Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR '14, pp. 737–746. ACM (2014)
366. Seamon, J.G., Moskowitz, T.N., Swan, A.E., Zhong, B., Golembeski, A., Liong, C., Narzikul, A.C., Sosan, O.A.: Sensecam reminiscence and action recall in memory-unimpaired people. Memory **22**(7), 861–866 (2014)
367. Sellen, A.J., Fogg, A., Aitken, M., Hodges, S., Rother, C., Wood, K.: Do life-logging technologies support memory for the past?: an experimental study using sensecam. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 81–90. ACM (2007)
368. Shen, W., Wang, J., Luo, P., Wang, M.: Linking named entities in tweets with knowledge base via user interest modeling. In: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '13, pp. 68–76 (2013)
369. Simond, F., Arvanitopoulos Darginis, N., Süsstrunk, S.: Image aesthetics depends on context. In: Proc. of Int. Conf. on Image Processing (ICIP). IEEE (2015)
370. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. CoRR, abs/1409.1556 (2014)
371. Singh, V., Elamvazuthi, I.: Fingerprint matching algorithm for poor quality images. The Journal of Engineering **1**(1) (2015)

372. Sinha, P., Mehrotra, S., Jain, R.: Summarization of personal photologs using multidimensional content and context. In: Proceedings of the 1st ACM International Conference on Multimedia Retrieval, ICMR '11, pp. 4:1–4:8. ACM (2011)
373. Skobeltsyn, G., Junqueira, F., Plachouras, V., Baeza-Yates, R.: Resin: A combination of results caching and index pruning for high-performance web search engines. In: Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '08, pp. 131–138. ACM, New York, NY, USA (2008). DOI 10.1145/1390334.1390359. URL http://doi.acm.org/10.1145/1390334.1390359
374. Smit, E., Van Der Hoeven, J., Giaretta, D.: Avoiding a digital dark age for data: why publishers should care about digital preservation. Learned Publishing **24**(1), 35–49 (2011)
375. Snoek, C., Fontijne, D., van de Sande, K.E., Stokman, H.e.a.: Qualcomm research and university of amsterdam at trecvid 2015: Recognizing concepts, objects, and events in video. In: Proc. of TRECVID 2015. NIST, USA (2015)
376. Snoek, C.G.M., Worring, M.: Concept-Based Video Retrieval. Foundations and Trends in Information Retrieval **2**(4), 215–322 (2009)
377. Sokal, R.R.: Classification: purposes, principles, progress, prospects. Science **185**(4157), 1115–1123 (1974)
378. Solli, M., Lenz, R.: Color emotions for image classification and retrieval. In: Conf. on Colour in Graphics, Imaging, and Vision, vol. 2008, pp. 367–371. Society for Imaging Science and Technology (2008)
379. Spyromitros-Xioufis, E., Papadopoulos, S., Kompatsiaris, I.Y., Tsoumakas, G., Vlahavas, I.: A comprehensive study over vlad and product quantization in large-scale image retrieval. IEEE Transactions on Multimedia **16**(6), 1713–1728 (2014)
380. St. Jacques, P.L., Levine, B.: Ageing and autobiographical memory for emotional and neutral events. Memory **15**(2), 129–144 (2007)
381. St. Jacques, P.L., Schacter, D.L.: Modifying memory: Selectively enhancing and updating personal memories for a museum tour by reactivating them. Psychological science **24**(4), 537–543 (2013)
382. Subagdja, B., Wang, W., Tan, A.H., Tan, Y.S., Teow, L.N.: Memory formation, consolidation, and forgetting in learning agents. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2, pp. 1007–1014. International Foundation for Autonomous Agents and Multiagent Systems (2012)
383. Sun, G., Chen, Y., Liu, X., Wu, E.: Adaptive multi-task learning for fine-grained categorization. In: Proc. of the IEEE Int. Conf. on Image Processing (ICIP 2015), pp. 996–1000 (2015)
384. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)
385. Talarico, J.M., Rubin, D.C.: Confidence, not consistency, characterizes flashbulb memories. Psychological Science **14**(5), 455–461 (2003)
386. Tang, X., Luo, W., Wang, X.: Content-based photo quality assessment. IEEE Transactions on Multimedia (TMM) **15**, 1930–1943 (2013)
387. Tarjan, R.: Depth-first search and linear graph algorithms. SIAM journal on computing **1**(2), 146–160 (1972)
388. Thimbleby, H.: Improving safety in medical devices and systems. In: Healthcare Informatics (ICHI), 2013 IEEE International Conference on, pp. 1–13. IEEE (2013)
389. Tong, H., Li, M., Zhang, H.J., He, J., Zhang, C.: Classification of digital photos taken by photographers or home users. In: Advances in Multimedia Information Processing-PCM 2004, pp. 198–205. Springer (2005)
390. Trad, M.R., Joly, A., Boujemaa, N.: Large scale visual-based event matching. In: Proceedings of the 1st ACM International Conference on Multimedia Retrieval, ICMR '11, pp. 53:1–53:7. ACM, New York, NY, USA (2011). DOI 10.1145/1991996.1992049. URL http://doi.acm.org/10.1145/1991996.1992049

391. Tran, N.K., Ceroni, A., Kanhabua, N., Niederée, C.: Back to the past: Supporting interpretations of forgotten stories by time-aware re-contextualization. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, WSDM '15, pp. 339–348 (2015)
392. Tran, N.K., Ceroni, A., Kanhabua, N., Niederée, C.: Time-travel translator: Automatically contextualizing news articles. In: Proceedings of the 24th International Conference on World Wide Web, WWW '15 Companion, pp. 247–250 (2015)
393. Tran, T., Nguyen, T.N.: Hedera: Scalable indexing and exploring entities in wikipedia revision history. In: Proceedings of the 2014 International Conference on Posters & Demonstrations Track - Volume 1272, ISWC-PD'14, pp. 297–300 (2014)
394. Tran, T., Tran, N.K., Teka Hadgu, A., Jschke, R.: Semantic annotation for microblog topics using wikipedia temporal information. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP) (2015)
395. Tran, T.A., Schwarz, S., Niederée, C., Maus, H., Kanhabua, N.: The forgotten needle in my collections: Task-aware ranking of documents in semantic information space. In: Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval, CHIIR '16, pp. 13–22. ACM, New York, NY, USA (2016). DOI 10.1145/2854946.2854971. URL http://doi.acm.org/10.1145/2854946.2854971
396. Tsagkias, M., de Rijke, M., Weerkamp, W.: Linking online news and social media. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, WSDM '11, pp. 565–574 (2011)
397. Tschiatschek, S., Iyer, R.K., Wei, H., Bilmes, J.A.: Learning mixtures of submodular functions for image collection summarization. In: Advances in Neural Information Processing Systems, pp. 1413–1421 (2014)
398. Tuan, T., Schwarz, S., Niederée, C., Maus, H., Kanhabua, N.: The forgotten needle in my collections: Task-aware ranking of documents in semantic information space. In: Proceedings of the 1st ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR-16), March 13-17, Chapel Hill, North Carolina, USA. ACM, ACM Press (2016)
399. Tulving, E.: Elements of episodic memory (1985)
400. Tulving, E.: Episodic memory: from mind to brain. Annual review of psychology **53**(1), 1–25 (2002)
401. Tulving, E., Schacter, D.L., McLachlan, D.R., Moscovitch, M.: Priming of semantic autobiographical knowledge: a case study of retrograde amnesia. Brain and cognition **8**(1), 3–20 (1988)
402. Tulving, E., et al.: Episodic and semantic memory. Organization of memory **1**, 381–403 (1972)
403. Tzelepis, C., Mavridaki, E., Mezaris, V., Patras, I.: Video aesthetic quality assessment using kernel support vector machine with isotropic gaussian sample uncertainty (ksvm-igsu). In: Image Processing (ICIP), 2016 IEEE International Conference on, pp. 2410–2414. IEEE (2016)
404. Tzelepis, C., Mezaris, V., Patras, I.: Linear maximum margin classifier for learning from uncertain data. arXiv preprint arXiv:1504.03892 (2015)
405. Tzelepis, C., Mezaris, V., Patras, I.: Video event detection using kernel support vector machine with isotropic gaussian sample uncertainty (ksvm-igsu). In: International Conference on Multimedia Modeling, pp. 3–15. Springer (2016)
406. Underwood, B.J.: Interference and forgetting. Psychological review **64**(1), 49 (1957)
407. Vaittinen, T., Kärkkäinen, T., Roimela, K.: Experience explorer: context-based browsing of personal media. In: Symposium on Human Interface, pp. 111–120. Springer (2011)
408. Van Der Meulen, M., Logie, R.H., Freer, Y., Sykes, C., McIntosh, N., Hunter, J.: When a graph is poorer than 100 words: A comparison of computerised natural language generation, human generated descriptions and graphical displays in neonatal intensive care. Applied Cognitive Psychology **24**(1), 77–89 (2010)
409. Van Opdenbosch, D., Steinbach, E.: Avlad: Optimizing the vlad image signature for specific feature descriptors. In: Multimedia (ISM), 2016 IEEE International Symposium on, pp. 545–550. IEEE (2016)

410. Vargas, P.A., Fernaeus, Y., Lim, M.Y., Enz, S., Ho, W.C., Jacobsson, M., Ayllet, R.: Advocating an ethical memory model for artificial companions from a human-centred perspective. AI & society **26**(4), 329–337 (2011)
411. Vargas, P.A., Freitas, A.A., Lim, M., Enz, S., Ho, W.C., Aylett, R.: Forgetting and generalisation in memory modelling for robot companions: a data mining approach. In: Proc. Human Memory for Artificial Agents Symposium (AISB 2010), De Montfort University, Leicester, UK (2010)
412. Vennelakanti, R., Madhvanath, S., Subramanian, A., Sowndararajan, A., David, A., Dey, P.: Pixene: Creating memories while sharing photos. In: Proceedings of the 14th ACM international conference on Multimodal interaction, pp. 59–60. ACM (2012)
413. Viola, P., Jones, M.J.: Robust real-time face detection. International Journal of Computer Vision (IJCV) **57** (2004)
414. Vliegendhart, R., Larson, M., Pouwelse, J.A.: Discovering user perceptions of semantic similarity in near-duplicate multimedia files. In: CrowdSearch, pp. 54–58. Citeseer (2012)
415. W3C: PROV-O: The PROV Ontology - Technical Report. http://www.w3.org/TR/prov-o/ (2013). Retrieved September 2017
416. Walber, T., Scherp, A., Staab, S.: Smart photo selection: Interpret gaze as personal interest. In: Proceedings of CHI '14 (2014)
417. Wang, J., Allebach, J.: Automatic assessment of online fashion shopping photo aesthetic quality. In: Proc. of Int. Conf. on Image Processing (ICIP), pp. 2915–2919. IEEE (2015)
418. Wang, L., Wang, H., Zhu, F.: Twin feature and similarity maximal matching for image retrieval. In: Proc. of the 5th Int. Conf. on Multimedia Retrieval, pp. 59–66. ACM (2015)
419. Wang, Y., Dai, Q., Feng, R., Jiang, Y.G.: Beauty is here: Evaluating aesthetics in videos using multimodal features and free training data. In: Proc. of the 21st ACM Int. Conf. on Multimedia, pp. 369–372. ACM (2013)
420. Wang, Y., Lin, Z., Shen, X., Mech, R., Miller, G., Cottrell, G.W.: Event-specific image importance. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4810–4819 (2016)
421. Weibel, S.: The dublin core: A simple content description model for electronic resources. Bulletin of the American Society for Information Science and Technology **24**(1), 9–11 (1997). DOI 10.1002/bult.70. URL http://dx.doi.org/10.1002/bult.70
422. Weiss, Y., Torralba, A., Fergus, R.: Spectral hashing. In: Advances in neural information processing systems, pp. 1753–1760 (2009)
423. Wheeler, M.A., Stuss, D.T., Tulving, E.: Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. Psychological bulletin **121**(3), 331 (1997)
424. White, R.W., Bennett, P.N., Dumais, S.T.: Predicting short-term interests using activity-based search context. In: Proceedings of the 19th ACM international conference on Information and knowledge management, CIKM '10, pp. 1009–1018 (2010). DOI 10.1145/1871437.1871565. URL http://doi.acm.org/10.1145/1871437.1871565
425. Williams, H., Conway, M., Baddeley, A.: The boundaries of episodic memories., pp. 589–616. Oxford University Press (2008)
426. Wolters, M.K., Niven, E., Logie, R.H.: The art of deleting snapshots. In: CHI'14 Extended Abstracts on Human Factors in Computing Systems, pp. 2521–2526. ACM (2014)
427. Wolters, M.K., Niven, E., Logie, R.H.: The art of deleting snapshots. In: CHI'14 Extended Abstracts on Human Factors in Computing Systems, pp. 2521–2526. ACM (2014)
428. Wolters, M.K., Niven, E., Runardotter, M., Gallo, F., Maus, H., Logie, R.H.: Personal photo preservation for the smartphone generation. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, pp. 1549–1554. ACM (2015)
429. Wolters, M.K., Niven, E., Runardotter, M., Gallo, F., Maus, H., Logie, R.H.: Personal photo preservation for the smartphone generation. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, Seoul, CHI 2015 Extended Abstracts, Republic of Korea, pp. 1549–1554 (2015). DOI 10.1145/2702613.2732793
430. Wu, X., Hauptmann, A.G., Ngo, C.W.: Practical elimination of near-duplicates from web video search. In: Proc. of the 15th Int. Conf. on Multimedia, pp. 218–227. ACM (2007)

431. Xie, L., Hong, R., Zhang, B., Tian, Q.: Image classification and retrieval are one. In: Proc. of the 5th Int. Conf. on Multimedia Retrieval, pp. 3–10. ACM (2015)
432. Yan, L., Zou, F., Guo, R., Gao, L., Zhou, K., Wang, C.: Feature aggregating hashing for image copy detection. World Wide Web **19**(2), 217–229 (2016)
433. Yang, C.Y., Yeh, H.H., Chen, C.S.: Video aesthetic quality assessment by combining semantically independent and dependent features. In: Proc. of Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), pp. 1165–1168. IEEE (2011)
434. Yao, L., Suryanarayan, P., Qiao, M., Wang, J.Z., Li, J.: Oscar: On-site composition and aesthetics feedback through exemplars for photographers. Int. Journal of Computer Vision **96**(3), 353–383 (2012)
435. Yeh, C.H., Ho, Y.C., Barsky, B.A., Ouhyoung, M.: Personalized photograph ranking and selection system. In: Proceedings of the 18th ACM International Conference on Multimedia, MM ’10, pp. 211–220. ACM (2010)
436. Yeh, H.H., Yang, C.Y., Lee, M.S., Chen, C.S.: Video aesthetic quality assessment by temporal integration of photo-and motion-based features. IEEE Transactions on Multimedia **15**(8), 1944–1957 (2013)
437. Yilmaz, E., Kanoulas, E., Aslam, J.A.: A simple and efficient sampling method for estimating ap and ndcg. In: 31st ACM SIGIR Int. Conf. on Research and Development in Information Retrieval, pp. 603–610. ACM, USA (2008)
438. Yosinski, J., Clune, J., Bengio, Y., Lipson, H.: How transferable are features in deep neural networks? CoRR **abs/1411.1792** (2014)
439. Yu, H., Deng, Z.H., Yang, Y., Xiong, T.: A joint optimization model for image summarization based on image content and tags. In: Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, AAAI’14, pp. 215–221. AAAI Press (2014)
440. Zacks, J.M., Speer, N.K., Swallow, K.M., Braver, T.S., Reynolds, J.R.: Event perception: a mind-brain perspective. Psychological bulletin **133**(2), 273 (2007)
441. Zacks, J.M., Swallow, K.M.: Event segmentation. Current Directions in Psychological Science **16**(2), 80–84 (2007)
442. Zagoruyko, S., Komodakis, N.: Learning to compare image patches via convolutional neural networks. In: Conf. on Computer Vision and Pattern Recognition (CVPR). IEEE (2015)
443. Zhang, J., Marszalek, M., Lazebnik, S., Schmid, C.: Local features and kernels for classification of texture and object categories: A comprehensive study. Int. Journal on Computer Vision **73**(2), 213–238 (2007)
444. Zhang, W., Ji, J., Zhu, J., Li, J., Xu, H., Zhang, B.: BitHash: An efficient bitwise locality sensitive hashing method with applications. Knowledge-Based Systems **97**, 40–47 (2016)
445. Zhang, Y., Callan, J., Minka, T.: Novelty and redundancy detection in adaptive filtering. In: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR ’02, pp. 81–88 (2002)
446. Zhao, W.L., Tan, S., Ngo, C.W.: Large-scale near-duplicate web video search: Challenge and opportunity. In: Int. Conf. on Multimedia and Expo (ICME), pp. 1624–1627. IEEE (2009)
447. Zhao, Y., Hong, R., Jiang, J.: Visual summarization of image collections by fast ransac. Neurocomputing **172**, 48–52 (2016)
448. Zhou, J., Chen, J., Ye, J.: Clustered multi-task learning via alternating structure optimization. Advances in Neural Information Processing Systems (NIPS 2011) (2011)
449. Zhou, W., Li, H., Lu, Y., Tian, Q.: Sift match verification by geometric coding for large-scale partial-duplicate web image search. Trans. on Multimedia Computing, Communications, and Applications (TOMM) **9**(1), 4 (2013)
450. Zhou, W., Li, H., Lu, Y., Tian, Q.: SIFT match verification by geometric coding for large-scale partial-duplicate web image search. Trans. on Multimedia Computing, Communications, and Applications (TOMM) **9**(1), 4 (2013)