Social Aspects of Photobooks: Improving Photobook Authoring from Large-Scale Multimedia Analysis

Philipp Sandhaus and Susanne Boll

Abstract With photo albums we aim to capture personal events such as weddings, vacations, and parties of family and friends. By arranging photo prints, captions and paper souvenirs such as tickets over the pages of a photobook we tell a story to capture and share our memories. The photo memories captured in such a photobook tell us much about the content and the relevance of the photos for the user. The way in which we select photos and arrange them in the photo album reveal a lot about the events, persons and places on the photos: captions describe content, closeness and arrangement of photos express relations between photos and their content and especially about the social relations of the author and the persons present in the album. Nowadays the process of photo album authoring has become digital, photos and texts can be arranged and laid out with the help of authoring tools in a digital photo album which can be printed as a physical photobook. In this chapter we present results of the analysis of a large repository of digitally mastered photobooks to learn about their social aspects. We explore to which degree a social aspect can be identified and how expressive and vivid different classes of photobooks are. The photobooks are anonymized, real world photobooks from customers of our industry partner CeWe Color. The knowledge gained from this social photobook analysis is meant both to better understand how people author their photobooks and to improve the automatic selection of and layout of photobooks.

1 Introduction

Since the broad availability of consumer photo cameras, photos have been a means to capture important moment of anyone's life and to preserve them for personal usage and to show them to others. However, having only simple prints or files on a computer in a digital is not sufficient for many. Thoughts and memories connected

P. Sandhaus (✉)
OFFIS – Institute for Information Science, Oldenburg, Germany
e-mail: sandhaus@offis.de

S. Boll
University of Oldenburg, Oldenburg, Germany
e-mail: susanne.boll@uni-oldenburg.de

to events can far better expressed by taking the raw photo content as a means to design semantically rich multimedia presentations. Besides Web Galleries, travelblogs and other possibilities a popular way to do this are photobooks. Even in the analogue days photobooks were used to structure and enrich memories by arranging pictures in an album and enriching it with additional information and content. In the digital age this has become even easier: Service providers like CeWe Color[1] enable users to design photobooks on a home PC and transfer these presentations into a physical product. The digital way provides the user with additional means: Photos are sorted, clustered, resized, cropped, rotated and the pages are decorated with textual annotations, background and content from different sources, e.g. a geographical map to illustrate a hiking trip.

The sources for the photobook contents are not only the users' own harddisks anymore. With the growing success of photo sharing platforms and photo sharing in social networks, photobooks are more and more getting a representation of the users' social interactions: Important photos not only sit in the users' personal accounts, but are spread widely over their social network [17]. This social aspect is also present in the users' photobooks: Typical events seem to incorporate people to a quite large degree, and photobooks are often a reflection of the authors' social life.

With CeWe Color as a project partner we have access to a large number of structural representations of photobooks as well as their pixel-based content. These representations allow for an analysis of photobooks on a large scale. By exploiting large numbers photobooks we can strive to understand the structure and types or categories of albums. Furthermore, we can learn what kind of photos actually find their way into the album.

In this chapter we will analyze and discuss the social nature of photobooks. We aim to find out if and how social aspects are reflected. Additionally, we analyze the expressiveness of photobooks, i.e. we explore how explicit different events are described and how emotional this is done. The third aspect we analyze is the vividness of photobooks, i.e. are photobooks designed very colorful and lively or rather uniform and factual.

The remainder of this chapter is organized as follows. We start by giving an overview over relevant related works and then introduce the data model and characteristics of our test set. For further analyses of the data set we first describe the development of semantic classifiers for photobook classes. These classifiers are used in the photobook analyses regarding socialness, expressiveness and vividness.

2 Related Work

Research in the context of personalized multimedia presentations has been an active topic in the multimedia community for over a decade now. However, most works are

---

[1]CeWe Color is Europe's leading photo finisher.

focusing on the generation or authoring part. For this several frameworks and systems such as the Cuypers Multimedia Transformation Engine [25] and the SampLe System [9] have been developed. Important file formats for the representation of multimedia presentations are the W3C standardized SMIL [28], SVG [27] or Flash [1] format. The analysis of such multimedia presentations has only been addressed rarely. However, some works have approached the problem of missing semantic annotation from another perspective by giving support for semantic annotation during the creation of these presentations. One example is the SemanticMM4U Framework [20] which adds semantics support to personal multimedia presentations. It also became clear that different kinds of semantics can occur in the context of multimedia and that semantics are not static but vary according to multiple dimensions, like time, place and person. One attempt to structure the different types of semantics in MIR is given in [21]. The emergent character of semantics in personalized multimedia content is described in [22].

On the other side, several works are aiming to better, how people manage, share and interact with their digital photographs. An early work from an anthropologist's perspective analyzing the use of home videos and photos was done in [6]. A more practical study was done by Frohlich et al. [10] with the goal to find requirements for photoware, which the authors define as technologies enabling photo sharing. Van House [24] has done some early works in studying the uses of digital photos in general. One of the most prominent means to share photos are online communities. However, some works have explicitly focused on studying the way people share photos in such communities. A recent study analyzes different factors which impact the sharing of photos in online communities [16]. As activities before sharing, but after capturing Kirk established the terminus photowork [12]. An older, but still relevant study aiming at finding out how people manage their digital photographs was carried out by Rodden [18] by observing and analyzing the management habits of 13 participants over a period of 6 months. Crabtree [7] has analyzed the way people naturally collaborate around photos and share collections of photographs. Bentley [2] has studied the similarities between consumer photo usage and music usage.

The interest in understanding the human usage of photos has also led to dedicated workshops clustering different aspects of photo usage. In [11] the boundaries between activities regarding preparation of photo for sharing (photowork) and actually using these photos to communicate with others (phototalk) are challenged by analyzing means of collaborative photowork. Another workshop [14] has focused on the analysis of co-located photo sharing activities. These are sharing activities which take place at the same place and at the same time, e.g. watching a photo slide show together with friends.

Looking at works dealing with the semantic understanding of photos in general we see this has been a prominent topic for several years now. Starting with pure content analysis [23] more recent works have a stronger focus at combinations with context-based methods [4, 15]. [8] and [13] give a good impression over approaches of the last years. With MetaXa [3] the authors of this article have developed a semantic photo analysis library and framework which combines content- and context-based analysis methods semantic image annotation.
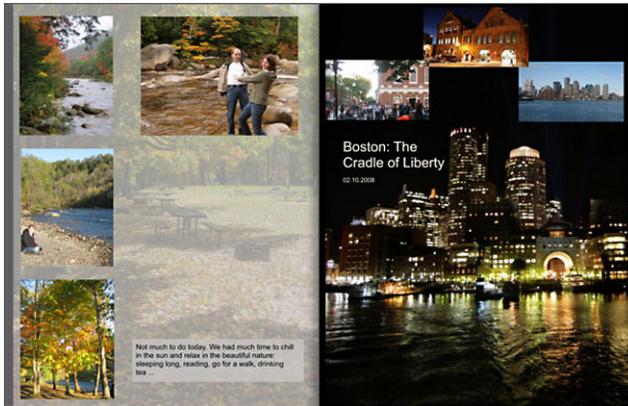
Fig. 1 Examples of typical photobook pages

3 Digital Photo Books

Personal photobooks have always been a popular way for organizing photos in a pleasant way and to preserve memories and share them with others. For this, people do not only place photos in a book, but decorate them with additional snippets like text annotations or page decorations. Have glue, scissors, and pencils been the tools in the analog days to assemble a photobook, the process has become digital nowadays. The user gets support by authoring tools like the one in [5] to digitally master photobooks. Such authoring software allows one to arrange digital images on the pages of an album, add textual annotations, and design the book with the preferred colors and style. In the end the user can order a print from a photo finishing company. Examples of photobook pages designed with such an authoring tool are shown in Fig. 1.

Creating a photobook is a form of digital story telling that reveals much about the user, the album and the different parts of the album. By authoring the photobook the user has implicitly enriched the photos, the single pages and the book as a whole with different kinds of semantics: She or he has established relations between photos, texts and pages. Photos and text become more prominent than others by their size and placement in the book, which may reflect their importance for the user. Some photos are clustered into groups or put on special pages, which might allow to draw conclusions about relations between photos and their semantics. We aim to reveal these hidden semantics by analyzing authored photobooks and the contained photos. Basis for this are the CeWe photobook software and photobooks which have been digitally authored and ordered.

To present the potential that photobooks offer for semantic analysis, we describe in the following the structure of photobooks and our data set.

Fig. 2 Photo Album Model (UML class diagram)

3.1 Photobook Structure

Photobooks are represented by a structural model capturing their content and layout over the pages of the book. This model serves as the input for the semantic analysis and directly represents the structural representation of photobooks ordered at CeWe Color. Figure 2 illustrates this structure in a UML model. It is divided into two parts, representing the photo album as a whole, the contained photos and their associated metadata.

A photo is defined by a number of features which stem from the photo's content and context together forming the *Photo Metadata*. These features are either directly extracted from the photo or derived by analyzing and combining other features.

Extracted features are, e.g., the capture time and location or the camera extracted from the photo's Exif header. Derived features are, e.g., the percentage of photos that contain faces or the brightness of the photo derived from the brightness histogram of the photo. A more thorough description of our framework for content-based and context-based metadata enhancement for photos is given in [3].

A *Photo Album* consists of one or more pages which each holds an arbitrary number of photos and text areas. A page can either be an album cover, the book spine or a regular page. It may hold a single image acting as the background. Photos and text areas can be of arbitrary dimension and can be rotated and overlap.

3.2 Test Data Set

Basis for our analysis are about 44,000 photobooks which have been ordered at CeWe Color. The photobooks have been authored and ordered by users with the help CeWe Photo Book software [5] in the time period from 3/2008 to 8/2010. From all orders in this time period, a portion of the daily orders of the day is selected and added to our test data set. The photobooks originated from all over Europe but the majority of orders come from Germany.

For our analysis, the photobooks have been completely anonymized, which means that all information about the person having placed the order has been removed from the photobook. Also, to maintain the privacy of the test data, our semantic analysis runs only on photo features, the photos themselves are removed from the data set. Thus, our test data set consists of all structural photo album information and the extracted photo features.

CeWe Color offers a wide variety of different kinds of photobooks ranging from low budget soft cover albums for a couple of Euros to premium books printed on photo paper in linen or leather cover. The photobooks of our test data set are not selected according to this criterion and therefore this variety is also reflected in our test data set.

4 Concept for Social Photobook Analysis

In this section we further explain the goals and concept of our social analysis of photobooks. Additionally we describe the development of semantic photobook classifiers to be able to distinguish different types of photobooks.

4.1 Goals

Our aim is to gain more knowledge about the social aspects of digitally authored photobooks. Specifically we are seeking to answer three questions.

4.1.1 How Social Are Photobooks?

We are interested in the question of how social aspects are represented in photobooks. Thus, how are social relations between people reflected in the photobook, how prominent are people shown in the book and are there differences for different types of photobooks.

4.1.2 How Expressive Are Photobooks?

By expressive we mean, how explicit is an event of a story documented in photobooks. A photo may, e.g., be not much more than a sole collection of photos, or it may be a detailed documentation of a specific event.

4.1.3 How Vivid Are Photobooks?

Photobooks can be either quite uniform and factual in their overall impression or more vivid and informal. One might suspect that this is different for different purposes or classes of photobooks, e.g. a book documenting a party might be more vivid than a professionally designed photobook of a wedding.

In the remainder of this chapter we will approach these questions.

4.2 Album Classification

One goal of our analyses is to detect differences between different types of photobooks. As shown in Sect. 3 these semantics are not directly present in the photobooks' structure of our test collection. Thus, we need a way to automatically extract these semantics for a large set of photobooks. In the following we describe the development of such classifiers by using a small portion of photobooks as a training data set.

The main requirement for the reliable derivation of characteristics for a specific semantic label is the availability of a sufficient large, labeled data set. This labeled data set is the ground truth for our analysis. Our problem is that we are equipped with a quite large quantity of data, but this data set is not semantically labeled in any sense. E.g., we do not know if a photobook is documenting a holiday or a wedding without looking at the book manually.

To compensate for this, we opted for a quite pragmatic method: We choose those samples from our test data set for which we are quite sure about their semantics according to the values of one or more features. By this we select only a small portion of the test data set as training samples. Samples of this labeled set are manually inspected to avoid wrong labels. With the help of these labeled samples we determine additional features discriminant for the respective semantic label. By this we

are able to derive rules for the semantic annotation of book parts, also for samples not fulfilling our initial characteristics. We are aware that this approach has two major drawbacks: The labeled training samples are not evenly distributed over the test data set and therefore we cannot be sure that derived rules for this labeled set can be used for labeling the rest of the data set. On the other side we cannot be sure whether or not the automatically determined ones do contain wrongly labeled samples. However, the manual inspection of samples of the automatically labeled data set determined by our approach has shown that our method is feasible.

From our and CeWe Color's experience we know that there exist some typical types of photobooks which are often ordered. Our goal was to determine features to distinguish between these different kinds of albums. For this we determined a set of labeled albums by selecting them by characteristics typical for specific kinds of photobooks. We then analyzed these labeled photobooks for additional discriminative characteristics.

4.2.1 Assumptions on the Data Set

The most typical albums according to CeWe Color are travel albums and albums documenting a party-like event, e.g. a birthday party or a wedding. Thus we opted to analyze our data set according to these typical events. Specifically, we chose three event types: a wedding and a birthday party and photobooks documenting a journey. We assume that we can select a considerably large subset of all albums belonging to these classes by looking for typical keywords in the title of the photo albums.

4.2.2 Ground-Truth Determination

We chose a quite pragmatic approach to select albums documenting such an event: We looked for typical keywords of on the title pages of the photo albums. Photobooks ordered at CeWe Color originate from countries all over Europe, but the majority of albums is ordered from Germany. We therefore restricted the input data set to albums ordered from Germany and selected albums that contained typically German keywords on the title page of the album. The selected keywords were *Hochzeit* (Wedding) for the wedding class, *Urlaub* (Holiday) for the travel class, and *Geburtstag* (Birthday) for the birthday class. We only looked for one very typical keyword for every class, as our goal was not to select as much albums as possible but to select the albums that quite reliably belonged to the respective class.

4.2.3 Analysis

Our goal was to find additional features which significantly differ when looking different classes of albums. For the classification of albums we opted to chose the following features:

	Wedding	Travel	Birthday
Average #Words	0.12	$< 10^{-4}$	0.09
Average #Images	0.004	$< 10^{-4}$	0.14
Auth. time	0.94	0.002	0.007
Time Span	0.28	$< 10^{-4}$	0.9
Average #Faces	$< 10^{-4}$	$< 10^{-4}$	0.0007

(a) Results of t-test
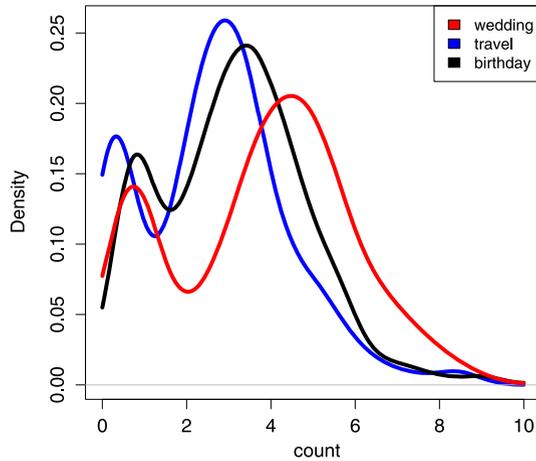
(b) Distribution of face counts

Fig. 3 Results (p-values) of Student's t-test for the significance of different features to decide about different album classes

- Average *number of words* per page
- Average *number of images* per page
- *Album authoring time*
- *Time Span* of the album
- Average *Number of Faces* per photo

We chose these features as we assumed that they would significantly differ for different classes. E.g., we assumed that a wedding album would show significantly more persons than a travel album.

Figure 3a shows the result of our analysis for the classification of albums. As in the determination of sub albums we performed a t-test for every album class and feature. A p-value lower than 0.05 designates a significant difference in the respective feature values.

4.2.4 Discussion and Classifier Training

One can see that most of the p-value are smaller than 0.05 which means that they are suitable to differentiate between the different classes of photobooks. The only feature that is significantly different in all classes is the average number of faces shown in a photo. Figure 3b shows the distribution for the respective classes. The average number of faces is 4 (wedding), 2.6 (travel), and 3.1 (birthday). The album time span is only significantly different for the travel class. This seems reasonable as journeys usually cover time periods of several days or weeks, while birthday and wedding events are only single day events. Another discriminative feature to distinguish between travel and other album classes is the number of words per page. We

observed that a travel album on average consists of 10.3 per page, while a wedding album consists of 6.3 and a birthday album of 6.5 words. This, again, seems reasonable as we observed that travel photobooks often contain rather long text passages in a diary-like manner. We used the labeled data set to train a Multiclass Naive Bayes classifier. The resulting multi-class classifier showed an accuracy of 79.46%.

We can conclude that, for a limited number of classes, we have identified features which quite reliably determine the type of a photobook. The results show that our approach is feasible and we can expect that the accuracy can be increased by tuning the parameters for the training of the classifier or by considering additional features. On the other hand we may expect that the accuracy of our classifier will decrease if we add additional classes of photobooks to it.

5 How Social Are Photo Books?

To answer the question if a photobook is also a *social* photobook or to which degree it is, we first have to define what makes a photobook social or how we can measure the *socialness* of a photobook. The common understanding of social is referred to as the kind and degree of interaction between two or more individuals. Thus, the socialness of a photobook can be described as the degree of social relations shown in the photobook. In addition, we can also differentiate these social relations on their degree of intimacy: How close are the relations of people in the photobook? Are they only strangers to each other or are they intimate friends or a couple?

We are aiming to at least to some extent find answers to these questions by analyzing the contents of the photobooks of our test collection. As indicators for the degree of socialness we have identified a number of indicators which will be further explored in the following.

5.1 Number of Persons

The presence of persons in a photobook is generally a good indicator that the photobook has a social aspect. The more people are shown, the more social the underlying event or story is. E.g. if a photobook shows many photos with people one may generally conclude that the photobook has a strong social aspect. However, if looking on a photo level, thus how many persons are shown in the photo, one may also be able to decide how much the people are related to each other.

In Fig. 4 the distribution of number of faces in a photo is shown. For this we employed the face detection algorithm proposed in [26]. We found out that over 85% of all photos in our test collection contain faces. The presence of persons in a photo is usually a sign that it is a quite personal and emotional photo. Knowing how many photos in an album contain faces can in turn tell us more about the type of the album. One may assume that a rather emotional event such as a wedding would also contain a lot of photos showing persons. Figure 4 shows the distribution over

**Fig. 4** Distribution of average number of faces per photo per photobook
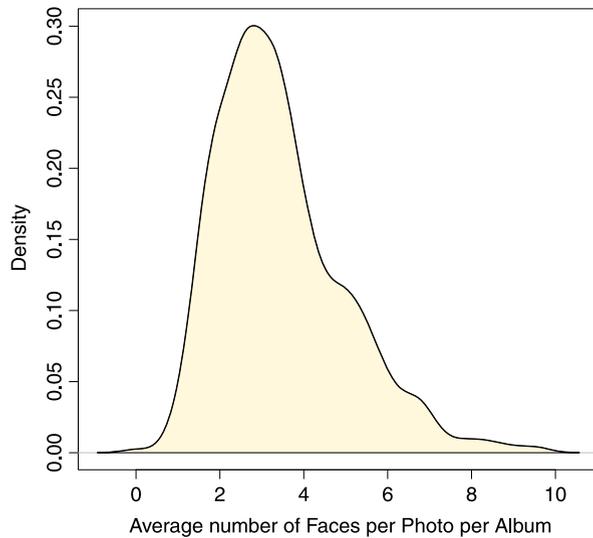
all photobooks for the average number of faces per photo. Obviously the majority of photobooks show an average of nearly two persons per photo.

Looking again at the results of our initial analysis of features for the development of a classifier for photobook types, we are also able to derive semantics for different kinds of photobooks. Figure 3b shows that the number of faces significantly differs for different photobook types: On average, travel photobooks show less people (2.6) in the contained photos than photos in birthday (3.1) or wedding photobooks (4). This backs up the common intuition that, e.g., weddings in general incorporate much more social aspects than e.g. a holiday trip. Also our manual inspection showed that almost all photos show people and often large groups of people e.g. in the wedding ceremony or at an evening party. On the other side, photobooks documenting a journey also show people, but to a much smaller extent, as not only the social relations are documented, but often the main topic are, e.g. a nice landscape or famous landmarks or buildings. Thus, we can conclude that there is a significant difference between different kinds of photobooks regarding the presence of social aspects.

5.2 Person Dominance

Not only the number of persons in a photo is an important indicator for the socialness of a photobook, but, perhaps more important, how dominant they are shown in the photo or the photobook. Thus, one person who occupies a large portion of a photo could indicate a portrait shot which would be considered as much more social than a photo which shows a large group of people, but only in the background.

Fig. 5 Face coverage and portrait shots
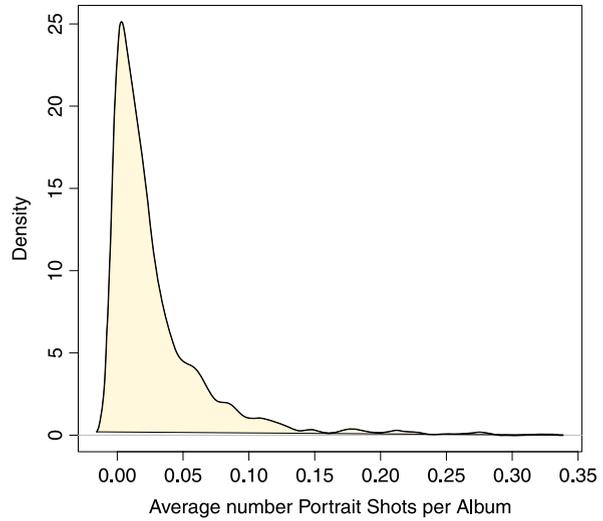
(a) Fraction of Portrait Shots of albums
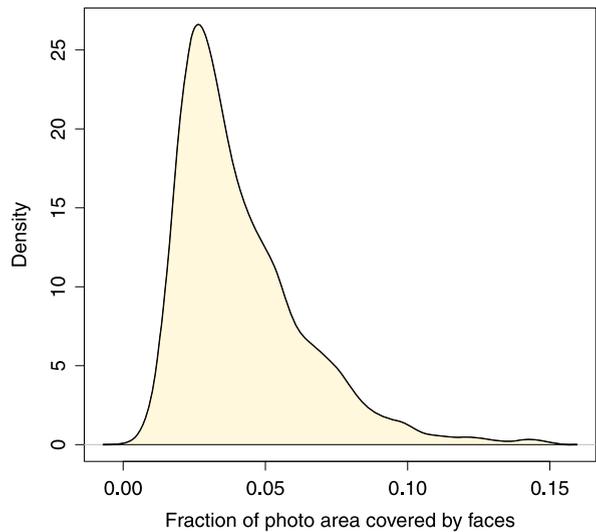
(b) Face Coverage

To cope with these differences we additionally determined the fraction of portrait shots taken over all photos in out test set. As portrait shots we have defined photos which show a single or two faces which occupy at least 20% of a photos' area. The result is depicted in Fig. 5a. The majority of photobooks does not have more than 10% of photos showing portraits. Giving the average number of faces per photo (nearly 2) per photo photobook, this seems to be not much, but backs up
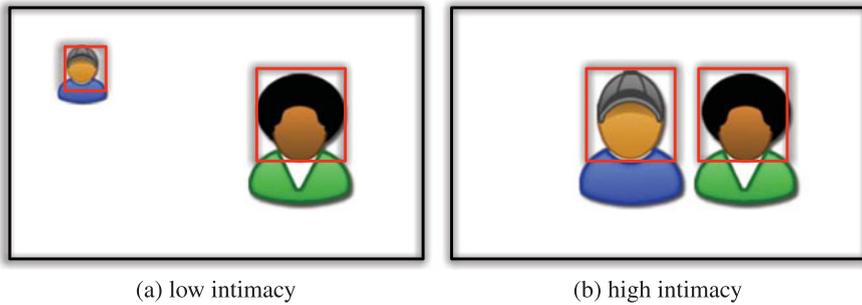
(a) low intimacy   (b) high intimacy

Fig. 6 Examples showing high and low levels of intimacy in a photos based on the size and closeness of shown faces

our visual inspection of the photobooks: Most photos showing people are showing their full body or additional objects besides the person itself. Typical examples are, e.g. photographs with one or two persons in front of a landscape or a famous building. Figure 5b shows an overview over the area in photos showing faces which are covered with faces.

5.3 Intimacy

Besides the dominance of persons in photos another interesting indicator for the socialness of photobooks is the degree of intimacy of people shown. The level of intimacy of people shown in a photo can partly be decided by their spatial closeness. This is especially true for only small groups of people (2–3). An example is shown in Fig. 6. Thus, the distance can be seen to be proportional to the level of intimacy in the photos. In most cases this is not true for large groups of people, e.g. group shots, as here people tend to gather to fit on the photo regardless of their intimacy to each other.

As depicted in Fig. 6 we see the distance of people in photographs as an indicator for their intimacy. Thus, if two persons are placed very near, we see this as an indicator for a quite intimate relation. We formally express this by the following formula:

$$intimacy(f1, f2) = \frac{area(f1) + area(f2)}{dist(f1, f2)} - \alpha \left| area(f1) - area(f2) \right|. \quad (1)$$

We set the distance of two faces in relation to their size in the photo: If two faces are rather small, then a large distance indicates a less intimate relation than the same distance for larger faces. Additionally, if one face is depicted a lot larger than the other this also indicates a large spatial distance i reality as one face is placed in the foreground while the smaller is placed in the background. This is expressed by the right part of the formula above.

Fig. 7 Degrees of intimacy
for different album classes
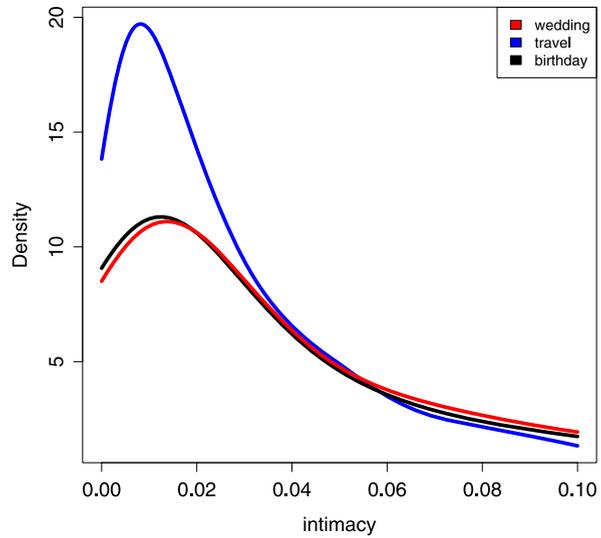
We have analyzed all photos of our data set showing two persons according to
this formula and have distinguished the photos regarding different classes of albums.
The result is depicted in Fig. 7. Interestingly, especially travel photobooks show a
high level of intimacy compared to wedding and birthday photobooks. A further
manual analysis of these travel photobooks showed that the photos showing two
persons are mainly photographs where two people took a shot of themselves in
a rather informal and intimate situation. Wedding photos with two persons were
comparably formal, e.g. taken at the reception of an evening party or formal photos
of the wedding couple.

5.4 Photo Origin

Besides at looking at the content of photobooks, one may also consider the origin of
the contained photos as an indicator of the socialness of a photobook. If the photos
in a photobook only originate from a single person, supposedly the author of the
photobook, this can be an indicator of lesser social involvement than if the photo
originate e.g. from a two or more people.

As a measure for the diversity of origins of photos in a photo the number of
different cameras can be taken. To a certain degree we can derive the number of
people having contributed photos to the photo album when assuming that every
person only owns one camera or one camera is not shared among a group of people
like a family. We know that this is a strong assumption, but from our own experience
and from interviewing people in our group we found out that people rarely own more
than one camera. The presence of more than one camera in an album therefore at
least seems to be a good indicator if more than one person was involved contributing

photos to the photo album. To determine the number of cameras we looked at the number of distinct values for the camera information in the photos' Exif headers of one album. To minimize the error we preprocessed the data set by removing photos without an Exif header or a missing value in the camera field, which, from our analysis of the data sets, is usually an indicator that it has been edited or scanned from a print.

The mean number of cameras over all photobooks in our test set is 2.8, which is a strong sign that not only one's own photos are used for an album but photos are shared among others. The photobook could be a compilation of photos from attendants of a holiday trip or the author of the photo could have added single photos from photo sharing sites or other sources to enrich his album.

6 How Expressive Are Photobooks?

Often photobooks are not only a means to preserve the memory to events for only the owner and creator, but can also act as vehicle to express one's feelings and thoughts to others. This can be done in various ways, either rather factual or more emotional. E.g. a photobook could have a very strict visual layout with none or only a few, but very precise and emotionless descriptions for the photos (like a photo of the Eiffel Tower with only the words *Eiffel Tower* as a description). A more emotional photobook would e.g. have a more casual layout (like tilted photos) and a more slang-like language (*Wow, what a cool view of that old tower!*).

In this section we are aiming at analyzing this expressiveness of real world photobooks. For this we will mainly focus on the text portions of photobooks.

6.1 Image–Text Ratio

In a first step, we analyze the ratio of text and image items in the photobooks of our test set.

Text Items per Page   The mean number of text items per page in our test set is 0.5. Only about 30% of all pages do contain text. This may be hint that only special pages are annotated with text and these pages may be a way to identify the beginning of a semantic unit in a photo album. Also, only 10% of the pages do contain more than one text item. This may be a hint that people tend to annotate pages as a whole rather than annotating single photos.

Images per Page   The mean number of photos per page is 2.9 and the majority of pages do not contain more than 3–4 images. This shows us that users, despite having the possibility to place much more photos on a page, prefer to have their photos being shown comparably large in a photo album.

Fig. 8 Distribution of font
sizes and text lengths over all
photobooks

(a) Distribution of font sizes

(b) Distribution the number of words per text item
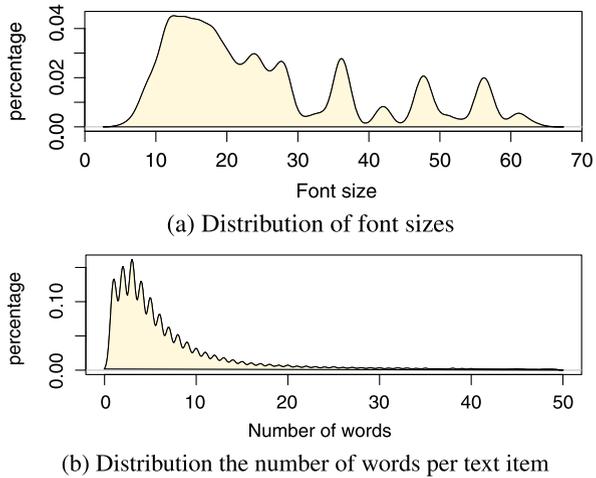
6.2 Text Lengths and Font Sizes

Besides photos the pages of an album can contain one ore more text items which
further describe one or more photos or one or more pages. We have already shown
that only about 30% of the pages contain text, which is a strong sign that such a
text description is a way to emphasize special parts of a photobook. One of the
main characteristics of text descriptions is their length in words and their font size.
Knowing these features may be a way to further semantically categorize text items
as, e.g., a page or image description. Thus we analyzed the text items in our test data
set regarding these two features.

The median number of words over all text items in our test set is five and 30% of
the text items contain up to three words. Thus, descriptions seem to be rather short
in albums and only very briefly give additional information to the album contents.
Text items with more words may designate a different kind of annotation, e.g. diary
type text describing a specific day or place of a holiday but not solely describing a
single photo. The distribution of text lengths is depicted in Fig. 8b.

In this context it may also help to take the font size of the text into account. We
found that the median text size is 22 and that 30% of the test items have a font
size of at most 16. Thus, these seem to be typical sizes for general annotations in
a photobook. A text with a font size significantly exceeding these values may be a
candidate to be a more important annotation such as an album title or the title for an
event in an album. The distribution of font sizes over all text items in our test set is
depicted in Fig. 8a.

7 How Vivid Are Digital Photobooks?

So far we have analyzed the degree of socialness and expressiveness of photo-
books. Another interesting characteristic is the vividness: Photobooks can be de-

signed rather factual and *cold* or more lively, e.g. by the use of many colors or strong variations in the visual layout. To some degree this is strongly related to the expressiveness of photobooks. However, unlike in the last section, we will focus more on the overall visual impression of photobooks rather than on the photobooks' textual contents.

As indicators for the vividness we have chosen two features, the diversity of colors and the intensity of the photobook pages. We have derived these features for the same photobook classes as in the last section.

7.1 Color Distribution

A way to rate the vividness of a photobook from a visual perspective is to determine how diverse the colors of the photos throughout one page are. E.g. a photobook with sepia or gray photos or photos which often show the same scene would be perceived as much less vivid than photos with very diverse colors. We have analyzed all pages of our photobooks regarding this diversity in colors and have compared this to different classes of photobooks with the help of our photobook classifier. For the comparison of photos we have employed the Color Distribution Descriptor defined in the MPEG-7 standard. For every album we have determined the average pairwise similarity of photos on a page.

The result is depicted in Fig. 9. Compared to travel and birthday photobooks, wedding photobooks show a much lower variation in their color layout throughout a page. This seems reasonable and backs up the impression from the manual, visual analysis of a number of wedding photobooks. These photobooks usually mainly incorporate very light and non-intensive colors and are limited to only a couple of colors. Thus, we can conclude that wedding albums are visually less vivid than other types of photobooks and create a rather calm, romantic and perhaps more intimate impression.

7.2 Intensity

As a second indicator for the vividness of photobooks we have identified the degree of smoothness or the intensity of the individual photobook pages. Thus we determined how many strong edges are present in the photos: A layout aiming at a more romantic style will often contain quite uniform areas with a strong use of depth-of-field smoothness. This effect is e.g. often used in professional portrait shots. Thus, we similarly analyzed our test set regarding the sharpness or contrast of contained photos and compared different classes of photobooks. For this we employed the MPEG-7 Edge-Histogram Descriptor and determined the percentage of non-edged pixels for every photobook page for different classes of photobooks.
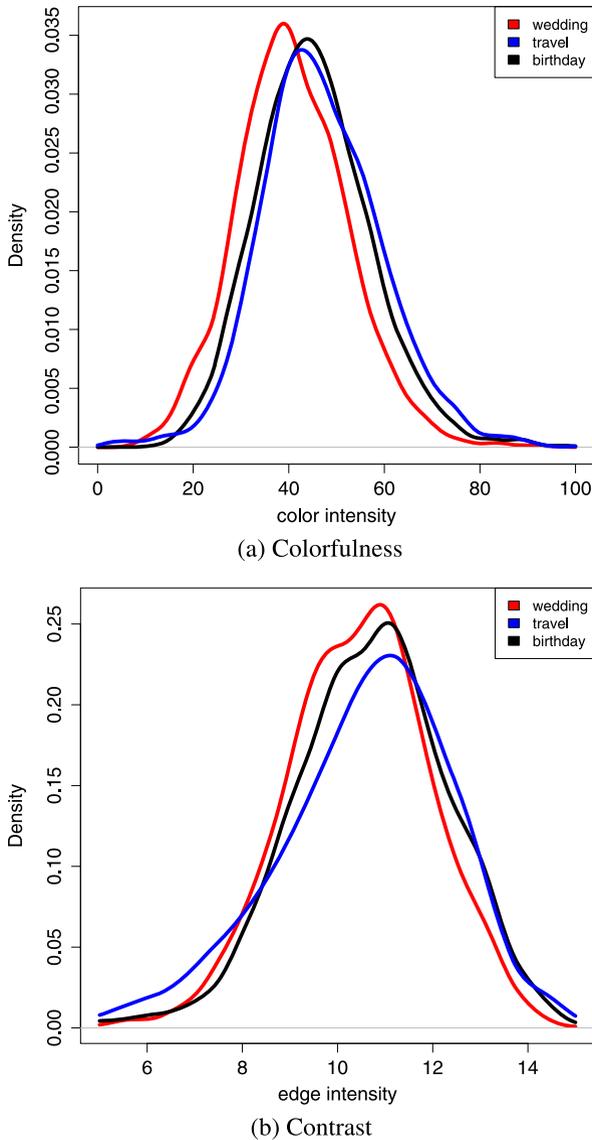
(a) Colorfulness

(b) Contrast

Fig. 9 Degrees of colorfulness and number of edges for different classes of photobooks

The result is depicted in Fig. 9b. Interestingly the degrees of edge intensity have a stronger variety in the travel photobooks compared to wedding or birthday photobooks. This can probably be explained by the variety of different events for travels. E.g. one could be on a winter skiing trip with a lot of smooth white areas in the photos or on a city trip where there are a lot of strong edges in the photos due to a lot of buildings. Also, again, the wedding photobooks shows significantly less strong

edges than the other photobook types, which also backs up our visual impression of wedding photobooks, which usually create a more intimate, less vivid and romantic impression, which is in line with the results of the color distribution analysis.

8 Conclusion

In this chapter we have presented the results of a large-scale analysis of real-world photobooks regarding the presence and degree of socialness, expressiveness and vividness. For this we have distinguished between different types of photobooks and have developed a classifier to automatically decide on the type of a photobook. In conclusion we can say that photobooks are showing very strong social aspects and persons are very prominent in most photobooks. However, we have also observed significant differences for different types of photobooks. One example is that wedding photobooks aim to create a much more intimate and calm atmosphere than other types. Although we limited the categorization to only three important photobook types, our analyses have shown that photobooks are a very interesting means to reveal more about peoples' incentives to design photobooks and how they express semantics and emotions for different types of events documented in these photobooks. We see the results in this chapter only as a first step. In the future we aim to identify additional photobook types and their special characteristics.

The main incentive for our analysis, besides to better understand peoples' behavior in photobook design, was to develop methods for automatic photobook design which take characteristics and implicit rules of real world photo books into account. The main conclusion of our analysis is that a meaningful layout system has to take into consideration the type of photobook which has to be created. E.g. different kinds of photos should be selected and a different layout should be chosen when designing a wedding photobook compared to a book documenting a journey. We will take this into account for future versions of our automatic photobook layout system [19].

References

1. Adobe: Flash file format specification 7. http://www.adobe.com/licensing/developer/ (2004)
2. Bentley, F., Metcalf, C., Harboe, G.: Personal vs. commercial content: the similarities between consumer use of photos and music. In: CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 667–676. ACM, New York (2006). http://doi.acm.org/10.1145/1124772.1124871
3. Boll, S., Sandhaus, P., Scherp, A., Thieme, S.: Metaxa—context- and content-driven metadata enhancement for personal photo books. In: Cham, T.J., Cai, J., Dorai, C., Rajan, D., Chua, T.S., Chia, L.T. (eds.) Advances in Multimedia Modeling. LNCS, vol. 4351, pp. 332–343. Springer, Berlin (2007)
4. Boutell, M.R., Luo, J.: Beyond pixels: Exploiting camera metadata for photo classification. Pattern Recognit. **38**(6), 935–946 (2005)
5. CeWe Color: CeWe Photobook (2007). http://www.my-cewe-photobook.com

6. Chalfen, R.: Snapshot Versions of Life. Bowling Green University Popular Press, Bowling Green (1987)
7. Crabtree, A., Rodden, T., Mariani, J.: Collaborating around collections: informing the continued development of photoware. In: CSCW '04: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, pp. 396–405. ACM, New York (2004). http://doi.acm.org/10.1145/1031607.1031673
8. Datta, R., Ge, W., Li, J., Wang, J.Z.: Image retrieval: Ideas, influences, and trends of the new age. ACM Comput. Surv. **40**, 51–60 (2008). http://wang.ist.psu.edu/survey/analysis/
9. Falkovych, K., Nack, F., van Ossenbruggen, J., Rutledge, L.: Sample: Towards a framework for system-supported multimedia authoring. In: Proceedings of 10th International Multimedia Modelling Conference, p. 362. IEEE, New York (2004)
10. Frohlich, D., Kuchinsky, A., Pering, C., Don, A., Ariss, S.: Requirements for photoware. In: CSCW '02: Proceedings of the 2002 ACM conference on Computer supported cooperative work, pp. 166–175. ACM Press, New York (2002). doi:10.1145/587078.587102. http://portal.acm.org/citation.cfm?id=587102
11. Frohlich, D., Wall, S., Kiddle, G.: Collaborative photowork: Challenging the boundaries between photowork and phototalk. In: Proc. of CHI Workshop on Collocated Social Practices Surrounding Photos (2008)
12. Kirk, D., Sellen, A., Rother, C., Wood, K.: Understanding photowork. In: CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 761–770. ACM Press, New York (2006). http://doi.acm.org/10.1145/1124772.1124885
13. Lew, M.S., Sebe, N., Djeraba, C., Jain, R.: Content-based multimedia information retrieval: State of the art and challenges. ACM Trans. Multimedia Comput. Commun. Appl. **2**(1), 1–19 (2006). doi:10.1145/1126004.1126005
14. Lindley, S.E., Durrant, A.C., Kirk, D.S., Taylor, A.S.: Collocated social practices surrounding photos. In: CHI '08: CHI '08 Extended Abstracts on Human Factors in Computing Systems, pp. 3921–3924. ACM, New York (2008). http://doi.acm.org/10.1145/1358628.1358957
15. Luo, J., Boutell, M., Brown, C.: Pictures are not taken in a vacuum. IEEE Signal Process. Mag. **23**(2), 101–114 (2006)
16. Nov, O., Naaman, M., Ye, C.: Motivational, structural and tenure factors that impact online community photo sharing. In: Third International Conference on Weblogs and Social Media (ICWSM), San Jose, California (2009)
17. Rabbath, M., Sandhaus, P., Boll, S.: Automatic creation of printable photo books out of your stories in social networks. In: 2nd SIGMM Workshop on Social Media (WSM2010), Florenze, Italy (2010)
18. Rodden, K., Wood, K.R.: How do people manage their digital photographs. In: Gilbert Cockton, P.K. (ed.) Proceedings of the Conference on Human Factors and Computing Systems, pp. 409–416. ACM, New York (2003)
19. Sandhaus, P., Rabbath, M., Erbis, I., Boll, S.: Employing aesthetic principles for automatic photo book layout. In: Proc. of International Conference on Multimedia Modelling, Taipei, Taiwan (2011)
20. Scherp, A.: Semantics support for personalized multimedia content. In: Proceedings of the IASTED International Conference on Internet and Multimedia Systems and Applications, pp. 57–65. ACTA Press, Innsbruck (2008)
21. Scherp, A., Jain, R.: Towards an ecosystem for semantics. In: MS '07: Workshop on Multimedia Information Retrieval on the Many Faces of Multimedia Semantics, pp. 3–12. ACM Press, New York (2007). http://doi.acm.org/10.1145/1290067.1290069
22. Scherp, A., Boll, S., Cremer, H.: Emergent semantics in personalized multimedia content. In: Fourth Special Workshop on Multimedia Semantics (WMS), Chania, Greece (2006)
23. Smeulders, A.W., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Mach. Intell. **22**(12), 1349–1380 (2000). http://doi.ieeecomputersociety.org/10.1109/34.895972
24. Van House, N., Davis, M., Takhteyev, Y., Ames, M., Finn, M.: The social uses of personal photography: Methods for projecting future imaging applications. University of California, Berkeley, Working Papers 3, 2005 (2004)

25. Van Ossenbruggen, J., Hardman, L., Geurts, J., Rutledge, L.: Towards a multimedia formatting vocabulary. In: Proceedings of the 12th International Conference on World Wide Web, pp. 384–393. ACM, New York (2003)
26. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (2001). citeseer.ist.psu.edu/article/viola01rapid.html
27. W3C: Svg. http://www.w3.org/TR/2004/WD-SVG12-20040510/ (2004)
28. W3C: Smil. http://www.w3.org/TR/2005/REC-SMIL2-20051213/ (2005)