MetaXa—Context- and content-driven metadata enhancement for personal photo books

Susanne Boll[1], Philipp Sandhaus[2] Ansgar Scherp[2], and Sabine Thieme[3]

[1] University of Oldenburg, Germany
Susanne.Boll@informatik.uni-oldenburg.de
[2] OFFIS - Institute for Information Technology, Oldenburg, Germany
{sandhaus|scherp}@offis.de
[3] CeWe Color AG & Co. OHG, Germany
thieme@cewecolor.com

Abstract. Making a photo book as a special gift to your beloved can be very time-consuming. One has to carefully select and arrange the pictures nicely over the pages of a previously bought photo book. In these days, photo finisher companies are able to directly print and bind a nice photo book from a selected set of images. But for the users of the software that comes with the album creation, the selection and arrangement of pictures in the album still remains a tedious task. What is missing are easy and good suggestions which pictures to select and how to arrange them into a personal photo book. A higher availability of metadata with the pictures could enable a content-driven and context-driven selection and make album creation better and easier. With MetaXa, we propose a flexible, component-based software architecture that iteratively allows for the multimodal extraction and enhancement of metadata for personal media content. The enhancement process is realized by extraction and enhancement components that each contribute to a well-defined annotation task. Depending on the application domain different components can be configured into a specific instance. With MetaXa, it is hence easy to reuse certain annotation algorithms in different scenarios and to alter a setup by adding or replacing certain enhancement components. MetaXa has been applied to the domain of photo book creation by our project partner CeWe Color and evaluated on a large set of consumer photos.

1 Introduction

Management and organization of one's personal photo collection is a laborious and therefore often never regarded task. Thousands of printed photos rest in the darkness and isolation of shoe boxes. In recent years, digital cameras became wide spread. However, they have not changed or solved the organization problem: Now it is pictures dsc2345.jpg to dsc2399.jpg residing in our digital shoebox, e. g., a folder called birthdayParty05. Currently, we are facing a market in which about 20 bn digital photos are taken per year for example in Europe [6]. At the same time, we can observe that many digital photos are never viewed nor used again. It is estimated that from all digital images only about 20% are actually printed [6]. This is not because we forget about these digital souvenirs. In a study [6] that our project partner CeWe Color, the world's leading photo finisher and digital photo service provider, has been carrying out by a market research institute is that most users of digital cameras would like to have their photos printed. Why are not more photos (re)used and printed even though it seems to be the customer's wish? The answer we give here is that the way in which we find and select photos from a large set of photos to print them needs far too much time and effort. The central insufficiency we face here today is the fact that digital photos today are just a poor reflection of the actual event captured. Digital cameras leave us with a pixel-based copy and some context information of what we experienced. Anything else is gone with the camera releaser, at least it is decoupled from the digital copy of the moment. Even though there is research in content-based image analysis for quite some years [25] as well as nice photo management tools [7, 1, 2], neither an automatic labeling nor a manual annotation of photos has become a success model.

What is needed is a better and more effective automatic annotation of digital photos that better reflects one's personal memory of the events captured. This approach would allow different applications to create value-added services on top of them such as the creation of a personal photo book. For this we propose a context-enhanced, multimodal method to achieve better image-understanding by the development of a novel, component-based software architecture.

Following the related work in Section 2, we present our content-based and context-driven metadata enhancement architecture (MetaXa), for iterative metadata extraction and enhancement of digital photos in Section 3. Based on our previous work on exploiting context for personal media collections [24, 3], we elaborate the design of the architecture and its components and present our multimodal metadata enhancement in Section 4. Section 5 describes the exploitation of the derived metadata in a concrete, professional photo book software to suggest a good (pre-)selection and composition of photos into an individual photo book, before we conclude the paper and present an outlook to future work.

2 Related work

In recent years, personal digital photo collections have received a great share of attention in the multimedia and database research community. In 2003, Rodden and Wood investigated if "advanced multimedia processing (speech recognition and content-based image retrieval) [are] useful in the context of personal photo collections?" [23]. The authors come to the conclusion that time and events are the preferred means of browsing through photo collections rather than advanced multimedia features. They also found out that manual annotation can not be expected form the everyday user. Interestingly, the participants still wanted to have prints. This observation very much complies with the study [6] that our project partner commissioned in 2003.

In the large and established research field of content-based image retrieval, we find approaches that address the domain of personal photo collections. In these approaches, content-based analysis, partly in combination with user relevance feedback, are used to annotate and organize personal photo collections. Prominent early systems are, e. g., MiAlbum [29], AutoAlbum [21], or SmartAlbum [27]. In the context of the DIVA project [19] a learning-based approach for content-based annotation of photos is introduced. In [30] content-analysis was used to automatically annotate photos based on face-recognition of family members. As the retrieved photos need to meet the users' expectation, a recent publication [14] proposes a hypothesis about human perception of image relevance. In the approaches referred to so far, the context of the photos is not included and exploited for the photo management and organization tasks. However, it became clear in content-based image retrieval that "One way to resolve the semantic gap comes from sources outside the image ..." [25].

With the availability of time and location from digital cameras, we find related work that aims to use this contextual information, sometimes in combination with content-based features, for organizing and accessing digital photo collections. In PhotoTOC [22], time and color histograms are used for the organization of photos in the visual user interface. Stating that "time matters" Mulhem et al. [15] define hierarchical temporal events as a clustering and organizational means. Also Graham et al. [8] consider "Time as Essence for Photo Browsing" in a calendar-based browser. Recently, FXPAL presented an elaborated temporal clustering for photo collections [5] based on similarity of time-stamps. In the ATLAS project, location and time are used for the organization of image collections [20]. Naaman et al. also exploit location for the automatic photo organization [17] and combine space and time for photo browsing [16]. In [4, 13] the authors discuss the use of content and context for scene classification.

Leaving the content-based field, there is also recent work in which only context information is used to annotate photos. In [18] identity-label suggestions are based on temporal, spatial, and social context. With the availability of EXIF header [11] for photos, this contextual information can be exploited for image understanding. The architecture presented in [12] proposes a context-based keyword creation for mobile video clips.

Considering the related work, we see approaches that either exploit content-based metadata, context-based metadata, or sometime also a multimodal combination of both to manage and organize photo collections. However, a systematic approach that combines content- and context-based metadata extraction and enhancement in the context of personal photo book applications does not yet exist. Consequently, we the propose with the MetaXa[4] architecture an approach that aims at embracing and advancing the state-of-the-art in multimodal metadata enhancement as well as systematically integrating content-based and context-based information. This approach is not only of academic interest but also commercially relevant as our concrete photo book scenario shows.

[4] Metaxa is a Greek liqueur, a blend of brandy and wine. Here, it means an architecture that is a blend of content- and context-based metadata enhancement.

3 Overview of the MetaXa architecture

The goal of MetaXa is to provide a component-based architecture for context-enhanced multimodal extraction and enhancement of semantic descriptions of personal photo collections. The architecture allows to configure the setup of metadata extraction and enhancement components. By the dynamic creation of an appropriate workflow a concrete instance of the architecture in realized to meet the specific metadata enhancement of a domain. This architecture not only allows to reuse components in different application domains but also to easily extend the architecture by new extraction and enhancement approaches.

3.1 The general MetaXa architecture

The central elements and features of the MetaXa architecture are illustrated in Figure 1. Input for the architecture are the photos taken by a digital camera. These photos, together with their contextual metadata, e. g., an EXIF header, enter the central *MetaXa Manager*. The image undergoes a sequence of extraction and enhancement steps, realized by separate components, in which metadata is enhanced iteratively. This allows for modularizing the metadata creation process into different steps. Increasing the amount and quality of available metadata, each step contributes to a better semantic description of the photos.

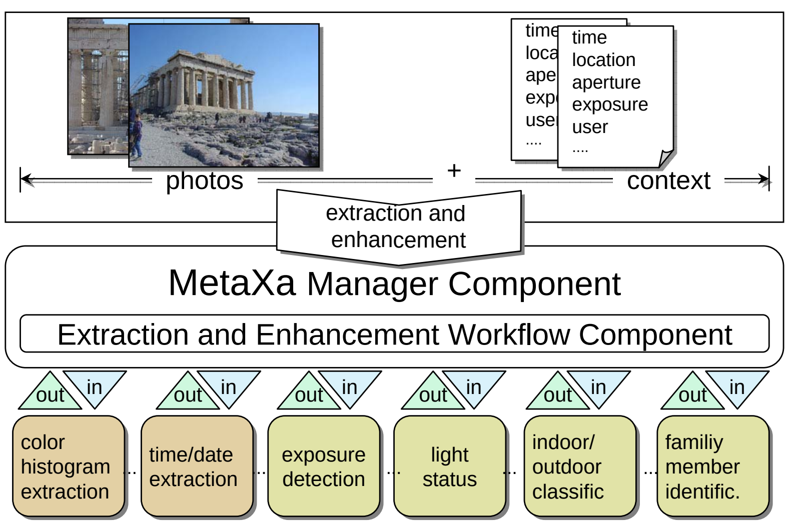


Fig. 1. Metadata extraction and enhancement architecture MetaXa

3.2 Metadata extraction and enhancement workflow

To ensure the correct order of enhancement the different extraction and enhancement steps are driven by a workflow. This workflow configures the sequence of steps that is carried out for each of the photos. The declarative workflow description identifies the different components and the manager component uses this workflow to drive the extraction and enhancement of the content. This allows to configure the architecture to the actual application needs without having to change the system.

For each extraction and enhancement component, an XML-file specifies which other metadata generated by other components are needed as prerequisites for their metadata extraction and enhancement. A dedicated workflow component is responsible for determining a workflow for this process on basis of the components' XML-specifications. This workflow ensures that all photos pass through the extraction and enhancement components in an reasonable order. This order is calculated by comparing the pre- and post-conditions of each extraction and enhancement component. The workflow component ensures that a specific enhancement component is only called if its pre-conditions are fulfilled. Since extraction components, i. e., a component which bases only on the raw photo data, don not have any pre-conditions, they can always be applied to a photo. The workflow component also detects circular dependencies between the components in order to prevent infinite loops.

3.3 Extraction and enhancement in MetaXa

In this section, we discuss the general design of the extraction and enhancement components of our system. Input to the architecture are the photos with their metadata. In a first phase, *extraction* components are used to extract relevant metadata and context directly from the photo. Hence, the system starts out with the information that comes with the digital photo itself, together with an optional EXIF header. Each photo is individually analyzed with different content and context feature extraction components that employ state-of-the-art methods. Examples of such components are color histogram extraction, edge detection but also the extraction of time, exposure time or GPS information from the EXIF header. Extraction components, illustrated in gray at the bottom of Figure 2, do not require previously generated metadata and hence form the starting point of the metadata enhancement.

On top of the extraction components, *enhancement* components derive new metadata which is stored for further enhancement and use. As illustrated by the edges between the components in Figure 2, enhancement components can use previously extracted or enhanced metadata to create new, potentially higher-level metadata. For this each enhancement component defines the necessary input which it requires to enhance the metadata. All components can still access the raw media content have access to all other photos in the collection, e. g., for similarity detection. Figure 2 illustrates examples of possible extraction and enhancement components and also indicates iterative use and enhancement of the

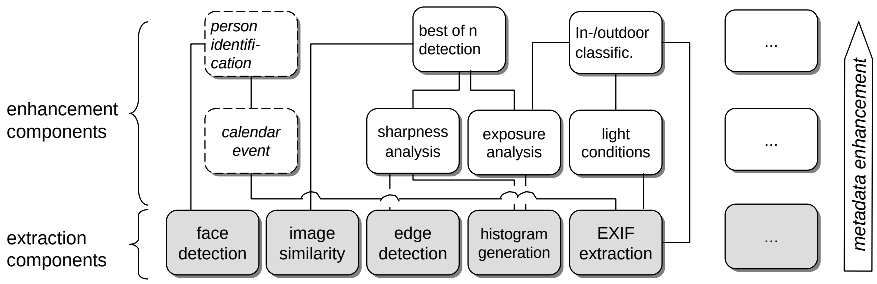


Fig. 2. Different types and dependencies of extraction and enhancement components

photos' metadata. Each component reveals its function to the MetaXa architecture. Depending on the application needs, our architecture allows to plug-in new components, such as the person identification and calendar event detection currently under development, and to configure the course and extent of extraction and enhancement.

4 Metadata extraction and enhancement components

Having presented an overview of the MetaXa architecture, we now describe the concrete design of the metadata extraction and enhancement components.

4.1 Design of extraction and enhancement components

The different extraction and enhancement tasks are realized by software components. Software components encapsulate their implementation and interact with the environment by means of well-defined interfaces [26]. Thus, a software component comes with a clear specification of what it *requires* and *provides*. As we employ Java to realize the MetaXa architecture, which doesn not natively allow for such a detailed specification, we use an XML-file for every component. These XML-files describe the kind and type of metadata that is necessary (pre-conditions) to apply a specific enhancement component. Since extraction components have no specific pre-conditions, they do not require a pre-condition description. However both extraction and enhancement components contain a section in this XML-file that specifies the provided metadata (post-conditions). The following listing shows an excerpt from such a file for the in-/outdoor enhancement component. This enhancement component requires the light status, time, brightness, and flash usage to determine whether the photo has been captured in-/outdoor.

```
<pre-conditions>
<metadatum description="time" relevance="mandatory">
  <mapping type="time-String" order="1">
    <source>Exif</source>
```

```
5   </mapping>
6   <mapping type="time-String" order="2">
7     <source>FileCreationTime</source>
8   </mapping></metadatum>
9   ...
10 <post-conditions>
11    <metadatum description="inoutdoor">
12     <mapping>In- or Outdoor</mapping>
13   </metadatum></post-conditions>
```

The example above shows that the component needs several previously generated metadata as input (indicated is only the metadatum time). The metadata entries can be marked as *mandatory* or *optional*. The same kind of metadata can be generated in different ways by different components, e. g., the information when a photo was actually taken. This could either be extracted from the Exif header of a photo or the file modification time. We meet this situation by the introduction of one ore more `mapping` entries in the XML-file. For example, the time metadatum can be achieved from the `source` Exif or file creation time (lines 3-8). A preference for a specific source (e. g., depending on the source's reliability) can be indicated by the parameter `order` (lines 3+6).

4.2 Extraction and enhancement components of MetaXa

For MetaXa we developed a set of concrete extraction and enhancement components of which we present representative examples. We developed components that exploit both image content and photo context. For this we employ state-of-the-art technology in content-based and context-based feature extraction and advance it toward context-enhanced multimodal photo metadata annotation.

Extracting content-based features. By content-based extraction we mean the purely pixel-based extraction of features of a photo. Typical components are histogram generation, edge detection, similarity analysis, and face detection. As examples, the similarity analysis and face detection are described in the following.

We developed a simple technique to determine *similarities* between photos. For it, we segment each photo into an 8×8-matrix and calculate the average RGB values. Each photo can then be described by three vectors, one for each RGB-channel. The similarity measure between two photos can then be described as the weighted sum of euclidean distances between the RGB vectors of the two photos. The human eye is not equally sensitive to the the colors red, green and blue. Thus, we weight the three color channels differently according to [9]. It is important to note that the size of the matrix has to be carefully chosen. If the segmentation is too coarse, the differences between two pictures can not be reliably detected as too many details get lost by calculation the averages, which can result in a false high similarity. However, if the segmentation is too fine grained, two pictures that, e. g., only differ in a small horizontal or vertical shift, would likely be considered very different. A good trade-off is a 8×8 matrix.

Face detection in the MetaXa architecture bases on the method presented in [28]. This method uses trained classifiers to rapidly detect faces in pictures. The classifiers were trained with several hundred positive samples and also a few hundred negative examples. The result is a *cascade of boosted classifiers*. The advantage of this algorithm is, that it is very fast and an open source implementation is available. We use this component to detect the number of faces in the image.

Extracting context-based features. These components utilize purely contextual information such as the EXIF header. *EXIF extraction* is used to extract the EXIF header information which is written to the photos by most consumer cameras. This header varies from camera to camera. However, most cameras at least provide information like timestamp, ISO, aperture, exposure, and if a flash was used. Some cameras also provide the focal length, the orientation and even location information such as a GPS position. Besides this, additionally we developed a component extracting general image information such as width, height, and time from the photo data in case there is no EXIF header available.

Content-based enhancement. Content-based enhancement components solely utilize information from content-based extraction- and other content-based enhancement components. As an example we present a component for sharpness analysis. *Sharpness analysis* is one of the key methods to determine the quality of an image. We developed two simple but fast methods: For both, we assume that most amateur photos contain a region of interest, which is located in the center. Consequently, we segment each photo in a 3×3-matrix and only analyze the center cell. The first method for sharpness detection uses the detection utilizing the Sobel filter [10]. We can use a resulting edge picture to determine an edge histogram. The photo is considered to be sharp if the histogram has high values in the upper bins. The second method utilizes the fact that in image regions, which are considered as sharp, often diverse levels of brightness occur. Thus, we use the brightness histogram and analyze how many bins exceed a certain value. The more of these bins exist the sharper the photo is considered to be. Both components provide this sharpness value to subsequent higher level metadata enhancement components.

Context-based enhancement. Context-based enhancement components solely utilize information from context-based extraction- and other context-based enhancement components. As example we present a component for determining the light conditions of a photo. *Light condition determination* is useful both for search purposes and as information for further analysis of the photos. Light condition can be derived from aperture and exposure time. For our component we employ the method described in [11] to calculate an exposure value from the given values in the EXIF header for aperture (F_n) and exposure time (E_t): $E_v = 2 * \log_2(F_n) - \log_2(E_t)$. The value E_v should be proportional to the brightness in the scene and therefore is a good indicator for the light condition.

Context-enhanced multimodal enhancement. These components utilize both context-based and content-based features for metadata enhancement. Here we present a component for classifying photos as in- or outdoor shots. *In-/Outdoor classification* is provided by a simple yet powerful method that relies on metadata extracted from content and context: the light conditions, daytime, if a flash was fired, and the exposure rating. The first two metadata entries are generated from context information and the last from the image content. For our in-/outdoor classification we apply a decision tree to the photos. This tree consists of rules that evaluate the before mentioned metadata. If, e. g., the picture is very dark, the flash was fired and it was taken at daytime there is a high probability that it is an indoor shot. A similar method has recently been proposed in [4]. Here Support Vector Machines are used to classify photos as indoor or outdoor. Unlike our approach only context information is used to classify the photos. In contrast, we aim at combining context and content information to achieve a higher precision. In a first step, we evaluated our approach with 437 consumer photos from which 68 are indoor and 369 are outdoor shots. This test set also compromises some ambiguous photos like indoor shots showing the view through a window. The component misclassified 0.5% of the outdoor shots as indoor and 10.1% of the indoor shots as outdoor. 10.5% of all pictures were classified as ambigous. These results are very promising, however, we are aiming at improving the classification accuracy by additionaly taking the EXIF header's object distance information into account.

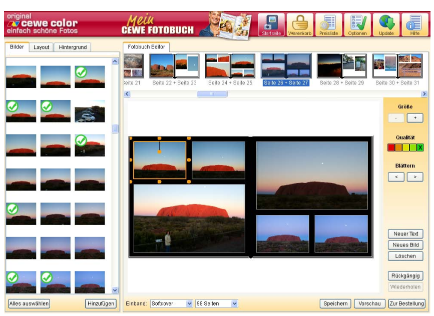
5 A concrete photo book application

Maybe one remembers having spent hours to create a nice photo album from a large set of of digital photos that carefully captures the impressions of a vacation or celebration: order prints, select the best pictures, sort and organize the photos along events and experiences, glue the photos into the album and label them – a tedious and time consuming task. In this section, we present the use and exploitation of the extracted and enhanced metadata within the personal photo book creation software of our project partner CeWe Color. An authoring wizard uses the extracted and enhanced metadata for the automatic "best-of" pre-selection of photos and layout of the photo book pages. The commercial software integrates the MetaXa results and the developed extraction and enhancement components.

The following Figures illustrate phases from photo book creation. Figure 3(a) shows an example of the different types of printed photo books the customer can chose from with CeWe Color's software. For an actual album creation, a user selects the photo collection from which the photo book should be created. Then the actual metadata extraction and enhancement of MetaXa takes place. This metadata is used for both a pre-selection of photos and their composition into a photo book. In a preferences dialog the users can influence this process by indicating which parameters should be taken into account for a "best-of" selection of photos by the software, e. g., sharpness, exposure, and similarity.

Figure 3(b) shows a photo book from a trip to Australia. Here, many similar photos especially of the Ayers Rock have been taken. Based on sharpness and similarity analysis, six photos are automatically chosen by the software for the photo book. These photos are shown in the center of Figure 3(b) and are in addition indicated by check marks on the left side. Figure 3(c) shows the results of an automatic selection of backgrounds. This is done, e.g., by taking the color histograms of the photos on a page into account. Having created a first version of the photo book, the users can still manually add, remove, resize, and rotate any photos and alter the backgrounds.

(a) page of a photo book



(b) preselected photos

(c) background selection

Fig. 3. Screenshots of the photo book application

The goal is to make the photo book creation an intuitive and easy task. The software[5] has been released in June 2006 and has been presented at the Photokina 2006 international trade fair in Cologne. Based on an evaluation of photo book orders, usability studies, and feedback from end users we will work on a refinement of the heuristics for pre-selection of photos and further extraction and enhancement components for MetaXa.

[5] http://www.cewe-photobook.com/

6 Conclusion

With the proposed MetaXa architecture, we presented an approach for the systematic integration of content-based and context-based metadata extraction and enhancement methods for digital photo collections. This architecture easily allows for instantiation of a specific setup of annotation methods to meet the requirements of a specific domain. It also allows to easily reuse components in different domains.

We employ and advance the state-of-the-art in semantic understanding of digital media in the domain of personal media. Our enhancement components exploit knowledge from content, context, and domain knowledge to provide for a better semantic understanding of photo collections and lay the grounds for next generation digital photo services.

The results of our MetaXa approach are evaluated by integrating them into the photo book software provided by CeWe Color. For it, we are evaluating and improving our extraction and enhancement components on large test sets provided by our photo finishing partner. We are also developing new extraction and enhancement components such as automatic orientation detection and location clustering.

On the application level, we are currently developing heuristics and probabilistic approaches for automatically suggesting relevant photos for a personal photo book. In addition, we are also working on dynamically determining a content- and context-based layout of the selected images in a photo book. Although the MetaXa architecture presented in this paper is applied in the domain of digital photo book authoring, it is not limited to this kind applications.

References

1. ACDSee Systems Int., Inc. acdsee Pro, 2006. http://www.acdsystems.com.
2. Apple Inc. iPhoto, 2006. http://www.apple.com/de/ilife/iphoto/.
3. Susanne Boll. Image and video retrieval from a user-centered mobile multimedia perspective. In *Proc. of the International Conference on Image and Video Retrieval (CIVR2005)*. Springer, 2005.
4. M. R. Boutell and J. Luo. Beyond pixels: Exploiting camera metadata for photo classification. *Pattern Recognition*, 38(6), 2005.
5. M. Cooper, J. Foote, A. Girgensohn, and L. Wilcox. Temporal event clustering for digital photo collections. *ACM Trans. Multim. Comp. Comm. Appl.*, 1(3), 2005.
6. GfK Group for CeWe Color. Usage behavior digital photography, 2006.
7. Google, Inc. Picasa, 2006. http://picasa.google.com/.
8. A. Graham, H. Garcia-Molina, A. Paepcke, and T. Winograd. Time as essence for photo browsing through personal digital libraries. In *Proc. of the 2nd ACM/IEEE-CS Joint Conf. on Digital Libraries*. ACM Press, 2002.
9. International Radio Consultative Committee (ITU). Encoding parameters of digital television for studios. Technical Report 601, CCIR, 1982.
10. Bernd Jähne. *Digital Image Processing*. Springer, 6th edition, 2006.
11. JEITA. Exif version 2.2. Technical report, April 2002.

12. Lahti, Westermann, Palola, Peltola, and Vildjiounaite. Context-aware mobile capture and sharing of video clips. In *Handbook of Research on Mobile Multimedia*. Idea Publishing, 2006.
13. Jiebo Luo, M. Boutell, and C. Brown. Pictures are not taken in a vacuum. *IEEE Signal Processing Magazine*, 23(2), 2006.
14. J. Martinet, Y. Chiaramella, and P. Mulhem. A model for weighting image objects in home photographs. In *Proc. of the 14th ACM Int. Conf. on Information and knowledge management*. ACM Press, 2005.
15. P. Mulhem and J.-H. Lim. Home photo retrieval: Time matters. In *Proc. of the 2nd Int. Conf. on Image and Video Retrieval*. Springer, 2003.
16. M. Naaman, S. Harada, Q.-Y. Wang, and A. Paepcke. Adventures in space and time: Browsing personal collections of geo-referenced digital photographs. Technical report, Standford University, InfoLab, 2004.
17. M. Naaman, Y. J. Song, A. Paepcke, and H. Garcia-Molina. Automatic organization for digital photographs with geographic coordinates. In *Proc. of the 4th ACM/IEEE-CS Joint Conf. on Digital Libraries*. ACM Press, 2004.
18. M. Naaman, R. B. Yeh, H. Garcia-Molina, and A. Paepcke. Leveraging context to resolve identity in photo albums. In *Proc. of the 5th ACM/IEEE-CS joint Conf. on Digital Libraries*. ACM Press, 2005.
19. W. K. L. P. Mulhem, J. H. Lim and M. Kankanhalli. *Advances in Digital Home Image Albums*, chapter 9. Idea Publishing, 2003.
20. A. Pigeau and M. Gelgon. Building and tracking hierarchical geographical & temporal partitions for image collection management on mobile devices. In *Proc. of the 13th annual ACM Int. Conf. on Multimedia*. ACM Press, 2005.
21. J. C. Platt. Autoalbum: Clustering digital photographs using probabilistic model merging. In *Proc. of the IEEE Workshop on Content-based Access of Image and Video Libraries*. IEEE Computer Society, 2000.
22. J. C. Platt, M. Czerwinski, and B. A. Field. Phototoc: Automatic clustering for browsing personal photographs. Technical report, Microsoft Research, 2002.
23. K. Rodden and K. R. Wood. How do people manage their digital photographs? In *Proc. of the SIGCHI Conf. on Human factors in comp. systems*. ACM, 2003.
24. Ansgar Scherp and Susanne Boll. Context-driven smart authoring of multimedia content with xsmart. In *Proc. of the 13th annual ACM Int. Conf. on Multimedia*, pages 802–803, New York, NY, USA, 2005. ACM Press.
25. A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 2000.
26. C. Szyperski, D. Gruntz, and S. Murer. *Component Software: Beyond Object-Oriented Programming*. Addison Wesley, 2nd edition, 2002.
27. T. Tan, J. Chen, P. Mulhem, and M. Kankanhalli. Smartalbum: a multi-modal photo annotation system. In *Proc. of the 10th ACM Int. Conf. on Multimedia*. ACM Press, 2002.
28. P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2001.
29. L. Wenyin, Y. Sun, and H. Zhang. Mialbum - a system for home photo management using the semi-automatic image annotation approach. In *Proc. of the 8th ACM Intl. Conf. on Multimedia*. ACM Press, 2000.
30. L. Zhang, L. Chen, M. Li, and H. Zhang. Automated annotation of human faces in family albums. In *Proc. of the 11th ACM Int. Conf. on Multimedia*. ACM, 2003.