Context-Aware Metadata Creation in a Heterogeneous Mobile Environment

Olga Volgin, Wanda Hung, Chris Vakili, Jason Flinn, and Kang G. Shin
Department of Electrical Engineering and Computer Science
The University of Michigan
{onv, wanda, cvakili, jflinn, kgshin}@umich.edu

ABSTRACT

With an exponentially-growing amount of digital information, data management is becoming increasingly burdensome for an average user. We propose an enhancement to the existing media-management techniques which utilizes the context available in the surrounding environment around the time a media file is created. This context information is associated with the file to provide enhanced data categorization and searching capabilities. A typical scenario considered in our design involves a heterogeneous environment where users capture digital images using camera-enabled mobile devices, such as iPAQs. Concurrently, these mobile agents also collect environmental data from a variety of sensors and other surrounding wirelessly-enabled devices. At the time of image capture, mobile devices collect and process environmental information, which is then associated with the digital image in the form of *metadata*. Association of the context with the image makes it possible to build user applications with context-based image classification and query facilities. Our system includes such a GUI application, which provides an image query mechanism based on metadata attributes collected at the time of a picture. Our preliminary evaluation of the system validates successful metadata creation and association with the media files and demonstrates the enhanced searching and classification capabilities using a GUI application.

Categories and Subject Descriptors

C.2.1 [Computer-Communication Network]: Network architecture and design; H.3.m [Information storage and retrieval]: Information Search and Retrieval

General Terms: Algorithm, Design, Human Factors

Keywords

Metadata, Heterogeneous Wireless Network, Digital Image Management

1. INTRODUCTION

The functionality of mobile devices is no longer restricted to communication or computation. Instead, mobile devices are often equipped with additional features including multimedia creation and processing capabilities that allow users to create and view digital data. Consequently, the amount of digital information accessible to users through the various mobile and stationary devices can be massive and makes the task of searching and organizing the data overwhelming.

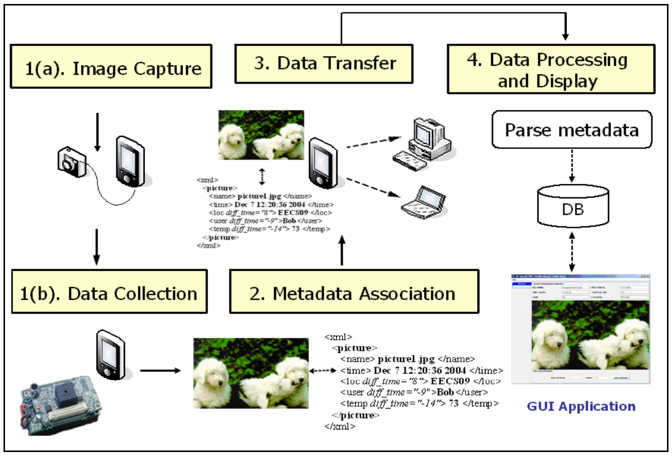
Currently, there is a limited number of content- or context-based consumer media management tools providing capabilities for organizing, searching, and browsing digital media [9]. Google Image Search [3] is a tool with the capability for text-based picture search, but its accuracy depends on the relevance of filenames and the text surrounding the images on the web-pages. The only existing alternative involves computation-intensive applications for analysis of digital media content. However, the approaches described do not utilize the semantic context at the time of image capture.

The main contribution of our work lies in its novel approach to utilize environmental information in context-based media-processing tools. In our system, environmental information is collected from wirelessly-enabled devices and the data is automatically associated with digital images. Our approach for data collection and association differs from existing techniques for context-based media systems and allows for a more meaningful digital media classification.

In our setting, a mobile user equipped with a camera-enabled device periodically captures digital images. The data available in the surroundings around the time of picture creation is most relevant and could include information such as user's location, current temperature, light, and the presence and/or identities of other people. Association of the metadata information with the image can be useful for providing tools for enhanced image query and management. For example, our approach allows a user to query for photos taken at a specific location on a sunny day. This information can be easily associated with pictures taken in the presence of environmental sensors. Our context-based approach and environmental data collection methodologies are unique because of the utilization of sensor nodes and other 802.11-enabled devices.

Once the images are created and associated with the relevant metadata, the images are transferred to a primary storage device, such as a laptop, where the metadata is parsed and stored as image attributes in a relational database. The GUI application provides the context-based image searching and management capabilities. An overview of the various stages in our system is presented in Figure 1, while the illustration of the mobile network and the environment envisioned is provided in Figure 2.

NOSSDAV'05, June 13–14, 2005, Stevenson, Washington, USA.



```
<xml>
  <picture>
    <name> picture1.jpg </name>
    <time> Dec 7 12:20:36 2004 </time>
    <loc diff_time="8"> EECS09 </loc>
    <user diff_time="-9">Bob</user>
    <temp diff_time="-14"> 73 </temp>
  </picture>
</xml>
```

Figure 1. System overview.

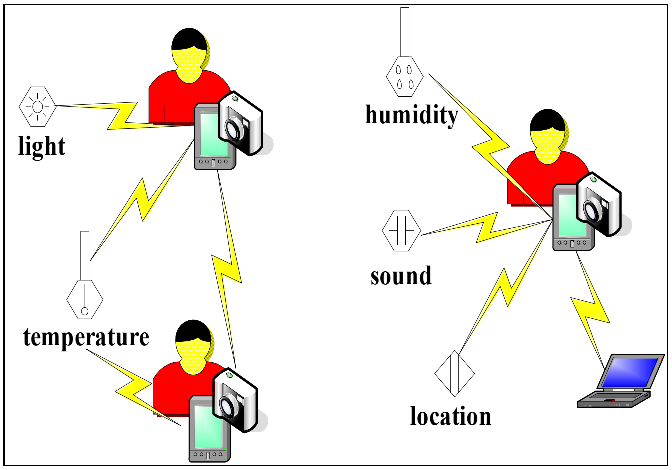


Figure 2. Mobile users acquire data from sensors.

As the technology moves forward, an increasing amount of information will become available in the environment due to the growing number of heterogeneous devices. A heterogeneous network in which mobile users acquire data from such sources could be beneficial for numerous purposes. Thus, modeling such systems is important and may lead to a discovery of numerous applications benefiting from this approach.

Our design offers multiple benefits. The first advantage is its flexibility in incorporating any wirelessly-enabled device capable of providing relevant information. Second, the use of XML for metadata creation allows for generic, yet expressive data representation while also enabling seamless addition of new metadata attributes. Although our system currently only considers digital image creation, this approach could be extended to provide metadata association for video/audio files. Furthermore, user-flexibility is incorporated into the system by allowing people to customize user preferences depending on their needs and the environment.

The rest of the paper is organized as follows. The related work is discussed in Section 2. Section 3 describes our design, followed by the system evaluation in Section 4. Section 5 provides concluding remarks and future directions.

2. RELATED WORK

Our approach builds upon the ideas presented in some of the related work in the area of heterogeneous networks and metadata collection, but it also introduces new insights. The idea of metadata creation for images captured with a cellular camera phone was first explored in [9]. This work relies on a server to perform image analysis and data classification while requiring human interaction to either accept or reject the classification. The system proposed in [11] requires manual metadata creation by the user at the point of image capture. Our design is preferred to these approaches because it eliminates the need for image processing while providing automatic metadata association based on environmental information and also permitting manual metadata association via the GUI application. The distributed nature of our approach improves reliability and eliminates the dependency on a centralized server.

The MyLifeBits system is designed to store and manage a lifetime's worth of data – everything that can be digitized [2, 8]. The advantage of our system is in its ability to provide media context by utilizing environmental data from third-party devices.

The authors of [1] utilize network configuration and the data-collection scheme closely resembling ours. However, the purpose for data collection is to supply a situational context to the mobile host device allowing it to modify its behavior in order to improve its service to a user. Our work applies a similar approach to context-based media metadata creation.

The authors in [6] also explore environmental detection with context-aware sensors. They explore *pull* and *push* communication schemes where *smart Context-Aware Packets* (sCAPs) capture information from different environmental sensors and allow devices to access the information. The Cricket system [10] illustrates a homogeneous network where a location can be determined and both active and passive architectures are explored. Our approach uses the pull (active) model to collect environmental data from a heterogeneous network, but our collection method differs from the one used in [6].

The automatic utilization of environmental information for context-based media classification distinguishes our approach from the existing work. The novelty of our design is reflected in the use of heterogeneous devices as the sources of context information for media files. The data-collection techniques are combined with the metadata-association mechanism to provide a context-based multi-media searching and classification mechanism that could be used independently, or combined with existing media processing applications.

3. DESIGN

This section discusses the overall design of our system and describes its various components which include network infrastructure, the metadata-association algorithm, and the GUI application for context-based query of media files.

3.1 Network Infrastructure

The primary network components in our system are e*nvironmental nodes*, *logical nodes*, and *mobile nodes*. The purpose of environmental nodes is to provide environment-specific data, such as light intensity, temperature, and location. In our system Mica2 sensor nodes (*motes*) [7] were used for this purpose. However, any other types or combination of sensor nodes, such as Intel Motes [5], could also be used. Although the motes used utilize a short range RF interface, sensor nodes with other interfaces such as Bluetooth and 802.11 could be incorporated.

Due to the heterogeneity of the interfaces of the sensor nodes and the 802.11-enabled mobile devices in our system, the sensor nodes were unable to directly communicate environmental information to the mobile devices. To overcome this problem, logical nodes were introduced. A logical node consists of a gateway mote connected to an 802.11 wireless-enabled laptop. A gateway mote attached to the laptop collects the data from the environmental nodes within range and forwards the information to the laptop.

Logical nodes serve two purposes. First, they act as a wireless 802.11 bridge between the environmental sensors and the mobile devices in our system. This setup allows RF-enabled sensors to communicate data to 802.11-enabled mobile nodes. This extra level of indirection between mobile and sensor nodes adds flexibility to our design by making it possible to incorporate sensor nodes with heterogeneous interfaces into our system. Additionally, the introduction of the logical nodes provides an energy-saving mechanism for the resource-challenged motes. This design is beneficial because it permits the sensor nodes to transmit the data to the logical node periodically and enter an energy-conserving state between those times. The logical node is a more powerful device connected to a power supply, and therefore, it can be used to store the data transmitted by the sensors and pass that information to mobile nodes on-demand.

The third major infrastructure component of our system is a mobile device, such as an iPAQ. Mobile devices collect the environmental data from the logical nodes as well as other 802.11-enabled devices and associate the data with media files based on temporal and spatial proximity. In order to accurately portray the scenario where mobile users roam in an environment while taking pictures, an ad-hoc mode of communication via UDP was used.

Since the focus of our work is not on the image creation, but rather metadata creation and its association with existing images, the image capture is simulated by user input. In an actual system, a digital camera is envisioned to be connected to, or built into the mobile device. When a picture is created the iPAQ requests data from devices in its wireless range. Our system uses an active architecture where the mobile node requests information while other devices passively monitor the network and respond to requests. Both logical nodes as well as other mobile devices reply with information. The logical nodes provide environmental information collected from sensor nodes while the information provided by mobile nodes is currently limited to the identification of the iPAQ's owner. Such information could be useful in identifying people in the picture. Privacy concerns, although currently not addressed, should be considered in future. The data received by the mobile node in response to its request is then processed to create metadata which is associated with the picture. This process is discussed in detail next.

3.2 Metadata Association

Our main contribution is the association of the environmental data with digital pictures captured by a user equipped with a mobile device. There are various approaches to the data-collection process. To maximize the amount of information available for association, our approach combines on-demand and periodic data collection.

On-demand data collection only considers the environmental data available at the time of picture creation. In this case, a request for data is sent out at picture capture time and the responses received are used for metadata creation. The periodic data collection occurs regardless of the user activity, and the information is stored for future association with media files as necessary. Both of these collection methods are used to optimize the amount and quality of information available at the time of image creation. On-demand data collection retrieves data most immediately relevant to the current image. On the other hand, periodic data collection maximizes the amount of information available for association since it spans the time interval before and after image creation. Periodic collection is most useful when the on-demand request is sent and there are no sensors in proximity to reply with data.

During the metadata association process, the data collected using both techniques is prioritized based on its time relative to the picture creation time. For maximum flexibility, the periodic collection time is controlled by the user's preferences. However, on-demand requests are issued regardless of the periodic collection time.

The process of environmental data analysis and its conversion to metadata begins at the instant a picture is taken. The amount of time before and after picture creation during which relevant data is collected is the metadata-association period. For example, a 30-second metadata-association period indicates that environmental data collected within 30 seconds before and after taking the picture is relevant. In this case, the metadata creation for the picture will be delayed by 30 seconds to collect sufficient data. The period is controlled by a predefined user setting.

The data collected is stored in the order in which it was received by the iPAQ. The information relevant to a particular picture is prioritized based on the offset time with respect to the time of picture creation. Data collected at smaller offsets has a higher priority due to increased perceived relevance. Additionally, duplicate data is filtered and thresholds are imposed to ignore small fluctuations in data caused by calibration and intrinsic hardware differences of the sensors. The collected data is processed to create metadata, which is associated with the picture in the form of an XML (eXtensible Markup Language) file. It should be noted that the self-describing nature and universal standards of XML ensure that our system can utilize any environmental data available and is not restricted to the items currently supported.

```
<xml>
  <picture>
   <name> picture1.jpg </name>
   <time> Dec 7 12:20:36 2004 </time>
   <loc diff_time="8"> EECS09 </loc>
   <user diff_time="-9">Bob</user>
   <temp diff_time="-14"> 73 </temp>
   <loc diff_time="-17"> EECS03 </loc>
  </picture>
  <picture>
   <name> picture2.jpg </name>
   <time> Dec 7 12:22:46 2004 </time>
   <loc diff_time="0"> EECS17 </loc>
  </picture>
</xml>
```

Figure 4. A representative metadata XML file.

During the metadata association process, picture filename and timestamp are mandatory while the environmental data is subject to availability. As shown in Figure 4, picture1.jpg file has information regarding temperature, user in proximity, light, and two location entries. The second file picture2.jpg has only a location entry.

Multiple locations for picture1.jpg are caused by movement of the user between the two locations during the metadata association period. The diff_time represents the positive or negative offset in seconds with respect to the time the picture was taken. In our system, the pictures and the corresponding XML files are stored on the iPAQ and are later transferred to a primary storage device, such as a laptop.

Our current implementation uses the cradle for data transfer from the iPAQ to a primary storage device. However, an alternative approach using wireless data transfer can be easily implemented. Once the data is transferred to the primary storage device, it is loaded into a GUI-based, database driven application. The application is used for managing, viewing, and searching media files. The following subsection discusses the database and the user application.

3.3 Database and User Application

The benefits of our approach can be observed through a GUI application enabling users to search and categorize images using context-based information. Our application parses and stores the information contained in the metadata to an MS Access relational database. It also provides an interface allowing users to search based on the information and to manipulate the metadata associated with images. Users may want to modify metadata or add additional information for a picture.

The application provides the end-user with the flexibility to search pictures based on attribute names and/or data values, and combinations thereof. Additionally, individual searches can be combined into more complicated queries with “OR” and “AND” operators. For example, a user may want to query for pictures taken at a certain location on a particular day when it was sunny. A user is also provided with a useful feature that allows searching based on a particular value even when he or she does not know the attribute name.

Once the user executes a query, pictures that match the criteria are returned. Any result can be selected and the available picture attributes are viewed in a sorted order relative to the time of the picture. Users are also able to modify existing attributes and add more attributes using the GUI to improve context accuracy. Additionally, we would like to add data-management capabilities allowing the user to organize files based on their queries.

The application described is implemented in Java with a backend relational database. Java was chosen due to its object-oriented style, flexibility, cross-platform portability, and available libraries. Also, Java allows web-accessibility, which allows users to manage their media from anywhere. Although the current application runs on a laptop, it could be ported to a mobile device providing greater flexibility even though mobile storage space is more constrained. We are currently exploring this possibility.

The application described is a sample query tool built to provide context-based data searching and management capabilities. Due to the well-known and generic format of XML, any number of custom data categorization tools could be built and used with our system, which is an advantage of our approach. Additionally, alternative application implementations utilizing XML databases could be considered as part of future work.

4. PRELIMINARY EVALUATION

The proposed system was evaluated on a testbed consisting of two 802.11b wirelessly-enabled h5555 iPAQs running Linux version 2.4.19 [4], thirteen environmental sensors, and three logical nodes. Each logical node was associated with various combinations of environmental sensors and created an area where mobile nodes could collect environmental data.

Two scenarios were tested in our environment. In the first case, the three logical nodes and sensors associated with them were spaced far enough to be outside of each others communication range. In the second case, two logical nodes were placed in proximity and the third logical node was placed remotely. In both scenarios, two iPAQ-equipped users roamed and simulated the action of taking pictures.

The system was evaluated based on multiple parameters. The first parameter was user-specified period, which determines the frequency of periodic data collection performed by the user’s mobile device. The metadata association period was the second parameter, which determined the window before and after picture creation containing the relevant data to be associated with the picture. The following subsections evaluate the system in areas of network communication, metadata association, environmental scenarios, and a comparison with closely related work.

4.1 Information Collection

Our system relies on communication between mobile clients and logical nodes for information collection. Clearly, an increased amount of environmental data leads to more useful metadata association. In our network infrastructure, requests initiated by a mobile node were addressed by the logical nodes, which responded with all available sensor measurements. Approximately three to four sensors were associated with each logical node and each data measurement was sent as a separate message. This configuration is reflected in the results, where the number of requests initiated by an iPAQ is three to four times fewer than the number of transmissions sent by logical nodes.

During our hour-long trial evaluation period, only two users at a time could explore the network due to the limited number of iPAQs. Each user performed a trial for 5 – 10 minutes and the iPAQ was then transferred to the next person. A total of nine users participated in the evaluation. The first seven users explored the first scenario of remotely-spaced logical sensors while the last two explored the alternate scenario. Table 1 lists the number of messages sent and received.

Table 1. Information loss between mobile nodes and sensors providing environmental information.

Data messages sent from logical nodes	1156
Data messages sent from mobile nodes	199
Number of received data messages for mobile nodes	1162
Message loss	**14%**

Messages were lost 14% of the time due to multiple causes. Some of the losses were attributed to rapid user movements in and out of range of logical nodes as well as to collisions over the wireless medium. The experiments took place in an indoor environment causing the structural obstructions to factor into losses. The results presented are representative of our experimental scenarios. However, different setups could yield different results. Although the message loss could be considered significant, it should be noted that the presence of cheap, redundant sensors could minimize some loss of information. Particularly, the message loss due to collisions could be insignificant because the data would still be available to the user through redundant devices.

4.2 Metadata Association

The experimental results indicate that the frequency of periodic data collection by the mobile device affects the accuracy of the metadata. Frequent periodic data collection makes it possible to collect the most data within the metadata creation window. However, the increased number of messages could lead to elevated energy consumption on the mobile device. We also observed that shorter metadata association intervals restrict the amount of available data while improving its relevance. Alternatively, longer association intervals allow more data to be collected and associated with the picture. Both the periodic data-collection interval and the metadata association period are user-controlled and can be selected depending on the environment.

4.3 Environmental Scenarios

As previously mentioned, two scenarios were examined in our evaluation. In one scenario the logical nodes were spaced far apart to avoid interference while in the second scenario two logical nodes were placed in proximity and the third node was out of communication range. These scenarios provided approximately the same results. In the case where logical nodes were in proximity, a lot of environmental data was reported by both logical nodes. This scenario ensured data redundancy, however, duplicated entries were filtered by the data-association algorithm. This scenario resulted in elevated energy consumption by the mobile node due to the increased number of messages. Remotely-spaced nodes provided the best results when larger area coverage was important.

4.4 Comparison with Related Work

The previous research that most closely resembles the proposed image-metadata association is reported in [9]. The evaluation presented there was based on user feedback through interviews and web surveys. The two major drawbacks of this approach are the limited availability and bandwidth of the cellular network, and the unfriendly phone interface required for user-interaction.

The limited availability is attributed to utilization of a remote server for file categorization, which makes the system unusable in areas of poor coverage. Also, user-interaction is required for every image, which could be burdensome. Our system increases the functionality of the mobile client through automatic metadata association, while not requiring user interaction.

In [9], the interface was difficult to use due to a small keypad and screen inherent to a cellular phone. Our system, on the other hand, allows users to conveniently view and manipulate data on a laptop or desktop. Our approach, however, does not allow for immediate user feedback, but it is not essential because the metadata association is automatic and editable via our GUI application.

Comparing these two systems is difficult. The cellular network used in [9] is much larger than our testing environment. For a fair comparison, a larger testbed with more powerful and numerous sensors is necessary and may provide interesting insights into the two systems.

5. CONCLUSION & FUTURE WORK

Data management can be an overwhelming task for mobile users as the amount of digital media information grows rapidly. In this paper, we proposed a novel approach to context-based metadata association for media files utilizing the information available in the surroundings. The association of environmental information with media files makes enhanced searching and categorization capabilities possible. Our design has several advantages. The use of heterogeneous devices including cheap, redundant, and ubiquitous sensors enables collection of context-based information. Furthermore, our approach does not require user input, yet it allows the users to modify context data as needed via the GUI application.

With an escalating amount of digital information and the lack of sufficient organizational tools, the metadata-association scheme presented in this paper could become increasingly important in the future. The potential future directions of this work include the evaluation of the system in a larger environment, possibly with more types of sensors and a real digital camera. Additionally, other communication interfaces, such as Bluetooth could also be evaluated.

6. ACKNOWLEDGMENTS

This work was supported in part by the US AFOSR under Grant No. F49620-01-1-0120.

We would like to thank Zhigang Chen and Min-Gyu Cho for their expertise and assistance with the iPAQs.

7. REFERENCES

[1] Gellersen, H. W., Schmidt, A., and Beigl, M. Multi-sensor Context-awareness in Mobile Devices and Smart Artifacts. *Mobile Networks and Applications*, vol. 7, no. 5 (October 2002), pp. 341–351.

[2] Gemmell, J., Lueder, R., and Bell, G. Living with a Lifetime Store. *ATR Workshop on Ubiquitous Experience Media* (September 2003), pp. 69–76.

[3] Google Image Search. http://images.google.com.

[4] Handhelds Documentation. http://www.handhelds.org.

[5] Intel Mote – Sensor Nets/RFID. http://www.intel.com/research/exploratory/motes.htm.

[6] Michahelles, F. and Samulowitz, M. Smart CAPs for Smart Its – Context Detection for Mobile Users. *Personal and Ubiquitous Computing*, vol. 6, no. 4 (September 2002), pp. 269–275.

[7] Mote Documentation and Development Information. http://www.cs.berkeley.edu/~awoo/smartdust/.

[8] MyLifeBits Project. http://www.mylifebits.com.

[9] Sarvas, R., Herrarte, E., Wilhelm, A., and Davis, M. Metadata Creation System for Mobile Images. In *Proc. of 2nd ACM/USENIX MobiSys* (Boston, MA, June 2004), pp. 36–38.

[10] Smith, A., Balakrishnan, H., Goraczko, M., and Priyantha, N. Tracking Moving Devices with the Cricket Location System. In *Proc. of 2nd ACM/USENIX MobiSys* (Boston, MA, June 2004), pp. 190–202.

[11] Wilhelm, A., Takhteyev, Y., Sarvas, R., Van House, N., and Davis, M. Photo Annotation on a Camera Phone. In *Extended Abstracts of CHI 2004* (Vienna, Austria, April 24-29, 2004), pp. 1403–1406.