Event-Centric View of Consumer Image Collections

Stacie Hibino and Mark D. Wood
Eastman Kodak Company
Rochester, NY, USA
stacie.hibino@kodak.com, mark.d.wood@kodak.com

Abstract— **Benchmark and ground truth databases are critical for evaluating imaging- and event-based algorithms, but such databases can be time consuming to create, cannot be representative of all real-life consumer image collections, and do not necessarily provide feedback on the general utility of an imperfect algorithm applied to real-life data. An alternative solution is needed to complement evaluation by benchmark and ground truth databases. We designed and developed Event Analyzer as a tool to meet this need. Event Analyzer combines data visualization of metadata with event and image retrieval results in a way that enables researchers to easily review data trends, quickly filter and select data, see relationships between metadata values, and visually scan retrieved results to determine correctness. In this paper, we describe the tool and discuss how it can and has been used to analyze real-life consumer collections in an event-centric manner.**

Keywords-metadata visualization; metadata validation; consumer image collection; events

I. INTRODUCTION

When new imaging- or event-based algorithms are developed to automatically extract semantic information or cluster images into semantic groups (e.g., face detection, event clustering), the accuracy and precision of such algorithms are typically measured against benchmark or ground truth databases of images. Although ground truth databases are critical for evaluating individual imaging- and event-based algorithms, they are plagued by three key issues: 1) creating a database of any substantial size is a time-consuming, manually intensive process, 2) it is impossible for any individual database to be representative of all possible real-life image collections, and 3) evaluating against such databases does not necessarily provide feedback on the general utility of an imperfect algorithm applied to real-life data. An alternative solution is needed to complement algorithm evaluations based upon ground truth and benchmark databases. In this paper, we present Event Analyzer (EA), a tool designed to meet this need.

The goal of EA is to provide researchers with a tool for assessing the general utility of several image- and event-based algorithms by reviewing their imperfect results applied to real-life consumer data sets and without requiring manual ground truth assignment of every possible metadata for every asset and event in a collection. We accomplish this goal by combining data visualization of metadata with event and image retrieval results in a way that enables researchers to easily review data trends, quickly filter and select data, see relationships between metadata values, and visually scan retrieved results to determine correctness.

Events are at the heart of EA and the tool provides an event-centric view of image collections. Previous studies show that consumers typically do not organize their image collections much; what little organization they do typically takes the form of creating date- and/or date and event-based folders at the file system level [9]. Although users are capturing more images more often, their largest clusters of photos are typically event-based. What once started as a few folders that could easily be browsed, has now become a multiyear collection spanning various holidays, life stages, and other events. As users' collections continue to grow, clustering becomes more important to aid them in managing and accessing their collections, and the concept of events is a natural clustering mechanism to leverage.

In this paper, we describe EA, discuss how different interactive data charts align well with consumer photo behavior and conceptions, show how EA can be used to examine the relationships between different metadata, and discuss how EA can be used to analyze real-life user data. The focus of this work is on *using* metadata derived from several algorithms and *not* the proposal of new algorithms. Thus, we do not describe details of algorithms for deriving metadata from images and events, but do provide overviews and references where appropriate.

Section II provides an overview of the metadata used by the system. In Section III, we describe EA. In Section IV, we provide examples of using EA with real-life data. In Section 0, we discuss related work and in Section VI, we provide a summary and discuss future work.

II. EVENT- AND ASSET-BASED METADATA

A. *Event Detection and Recognition*

Kodak's event detection algorithm groups images and video clips from a consumer's collection into a list of hierarchical events [6]. Event hierarchies consist of non-leaf event nodes and leaf nodes that represent media items (i.e., images or video clips) that were captured during a given parent event. The event hierarchy includes three levels of events: Super-Events, Events, and Sub-Events.

Event-level clusters are formed based upon analyzing capture date and time differences between media. If appropriate, Events are divided into Sub-Events using a metric of image similarity. Super-Events are formed by joining adjacent Events when the time differences between nodes have met a certain threshold. A consumer's collection can be summarized by "top-level" events, which we define as events that do not have an event node parent. Thus, a top-level event can be either an Event with no Super-Event parent or a Super-Event.

Kodak's event recognition algorithm [1] currently classifies each event into one of four common consumer event categories: party, sports, vacation, or family moments. This algorithm is a probabilistic classifier that uses event- and asset-level metadata such as inter-event time, time of day, indoor/outdoor, scene type, and number of detected faces.

B. Asset-Based Metadata

Metadata from an individual asset can be automatically extracted (e.g., from EXIF) or algorithmically derived. EA currently uses the asset-level metadata displayed in Table 1.

TABLE I. ASSET-LEVEL METADATA USED IN EVENT ANALYSIS TOOL

Metadata	Description
IVI Technical	Image Value Index from 1 to 4, indicating the algorithmically derived technical rating of an image [7]
Face Count	Number of faces included in an image as determined by a face detection algorithm
Focal Length	Focal length from EXIF
Exposure Time	Exposure time from EXIF
Flash	Flash value from EXIF
Outdoor Probability	Probability that an image was captured outdoors as determined by an outdoor scene analysis algorithm
Scene Classification	Numeric score representing the likelihood that an image includes a certain scene type [8]
Creation Date & Time	Capture date and time from EXIF
Geospatial Data	Latitude and longitude from EXIF (either from a GPS recorder or manually added)

C. Event-Based Metadata

Event-based metadata can be determined by the event detection or recognition algorithms, aggregating information from the assets belonging to an event or via new algorithms based upon asset-level metadata. EA currently uses the event-based metadata displayed in Table 2.

The event type is the highest ranked classification as returned by the event classifier algorithm [1]. The holiday keyword for an event is based on the automatically determined holiday labels for images belonging to that event. Multiple holidays may be associated with an image; a score is assigned to each holiday based upon how close the capture date is to the date of the holiday. Holiday scores for images are averaged together and EA assigns the highest scoring holiday keyword to the event.

Geospatial keywords include two types of information: place names and feature types. Both pieces of information are obtained by reverse geocoding the latitude and longitude of where a picture was captured, if available. The system uses the publicly available geonames.org web service [3] for the reverse geocoding. An event may span multiple geographic places; if so, EA uses the most frequently occurring place as the place of the event.

TABLE II. EVENT-LEVEL METADATA USED IN EVENT ANALYSIS TOOL

Metadata	Description
Media Count	Number of media contained in event
Date Range Start	Start date and time of event, based upon earliest date and time of media contained in event
Event Level	0 = Super Event, 1 = Event, 2 = Sub-Event
Event Type	Classification result of event recognition algorithm [1]
Event Hierarchy	Relationship between a selected event and any ancestor or descendent events
Holiday	The name of a holiday on which an event occurred, if applicable
Geospatial Keywords	The presumed GIS feature type(s) and place names associated with the location of the event

Feature types include terms such as "park," "mountain," and "school." In order to associate a feature type with a latitude and longitude, the system queries the geonames.org web service for the nearest feature points within a three-kilometer radius. It then uses a heuristic described in [9] to associate one or more of these feature types with the given latitude and longitude. EA uses the highest ranked feature type as the feature type. The feature types are defined by the geonames.org service, drawing upon other sources, including the U.S. Geological Survey's Geographic Names Information System definitions.

III. EA: ANALYZING EVENT METADATA

A. System Implementation and Overview

A series of extractors and algorithms are run to obtain and store the metadata described in Section II into an AllegroGraph triple store. This metadata extraction step only needs to be run once per each time that new media are imported into the system.

EA is a .Net 3.5 C# desktop application that accesses data from an AllegroGraph triple store. Queries to the triple store are posed using AllegroGraph Prolog. Since metadata is stored in a triple store format, it is simple to add new metadata from new, updated, or competitive algorithms, and access this new metadata from within EA.

A screen shot of the EA application is presented in Figure 1. The application window is divided into three columns: collection- and event-based metadata, media display, and asset-based metadata. We discuss and describe the details of EA in the remainder of this section.

B. Collection- and Event-Based Metadata

The leftmost column provides a summary of the currently loaded collection along with details of all top-level events contained within the collection. For example, in Figure 1. the collection summary indicates that the currently viewed set contains 2153 assets and 240 top-level events. Metadata details about top-level events are displayed by column in the data list text box. These details include the following metadata described in Section II.A: (total leaf) media count, date range start, event level, event type, outdoor probability, holiday, holiday probability, place type, place type score, city, and city score. Users can sort the text list of events by simply clicking the column heading by which

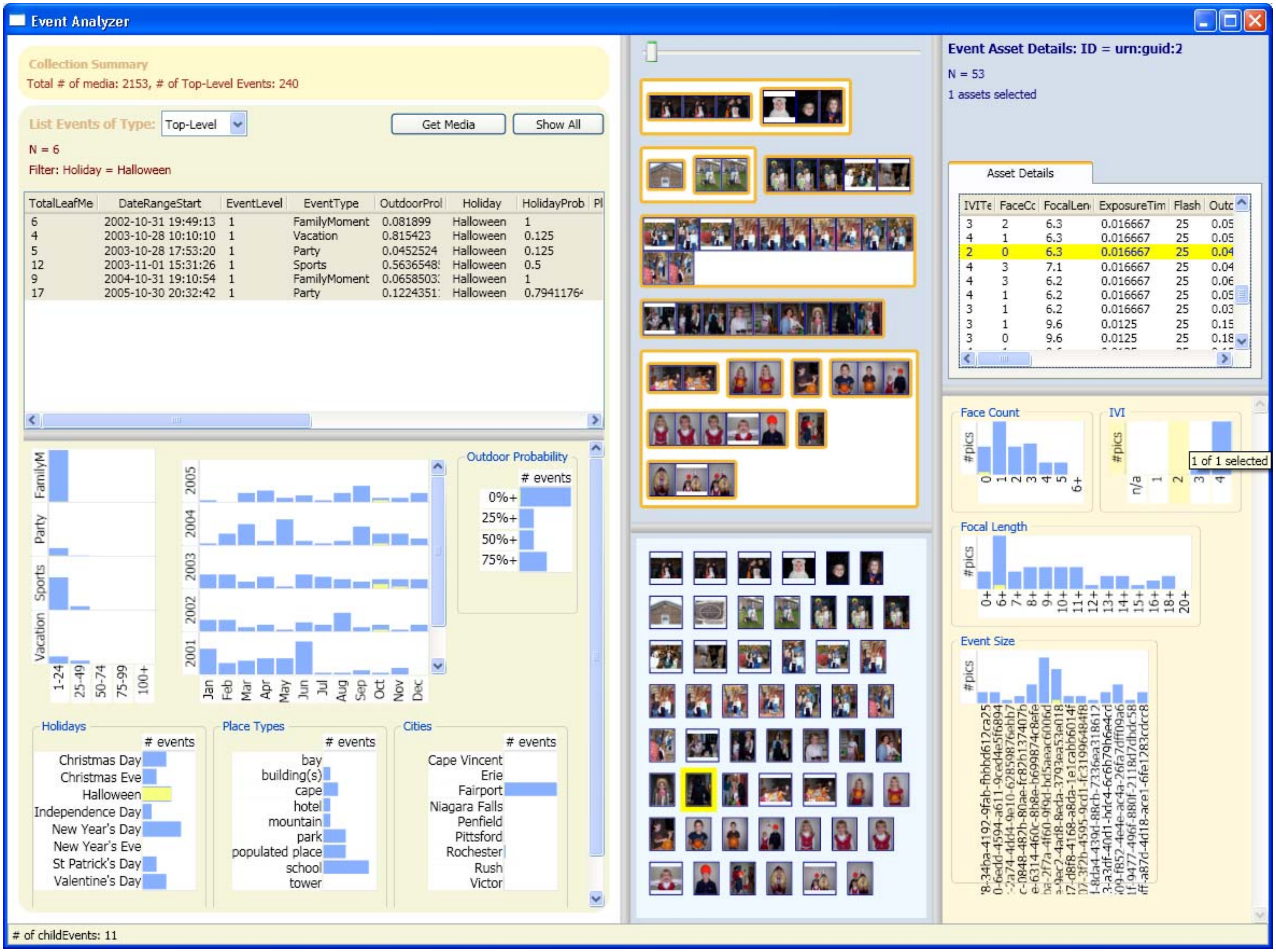

Figure 1. Sample screen shot of the Event Analyzer (EA) application.

they wish to sort. Clicking a heading once will forward-sort the list by the corresponding data column; clicking a heading a second time will reverse-sort the list by that column. The event list displayed in Figure 1. has been sorted by DateRangeStart, the starting date and time of the events.

Users can filter the text list of events by using the interactive bar charts displayed in the bottom portion of the event details area (see charts in bottom left area of Figure 1.). Any chart can be used to filter the list of events. The user simply clicks on a chart label on the x- or y-axis or on a bar within a chart to select events with the corresponding values. Blue bars represent the distribution of events for the whole collection; yellow highlighting within bars indicates the proportion of the bar currently selected. As is common in interactive data visualizations, selection is linked across charts so that a selection made in one chart results in corresponding selection highlighting in all other charts.

For example, Figure 1. shows the case where a user has selected the "Halloween" filter and that six events from the collection with some probability of being related to Halloween are displayed in the event text list. The corresponding October bars are highlighted in the month-year chart, along with part of the bar representing November 2003, since one event from November 1, 2003 has the potential of also being a Halloween-related event. Although less obvious, the other charts also have partial highlighting. In the situation where a bar is short and values or highlighting may be difficult to determine, users can simply hover with the mouse cursor over a bar area to receive bar selection and total information, as indicated in the IVI chart in the right column of Figure 1. Furthermore, hovering with the mouse cursor over an x- or y-axis label summarizes the counts for all bars with the corresponding values. For example, hovering the mouse cursor over the year "2002" label would summarize how many events occurred in 2002.

The event-based metadata columns and charts included in EA were chosen based upon consumer photo behavior along with how consumers think about their personal image collections. As mentioned in the introduction, consumers often think of their collection in terms of dates and events [9]. In addition to date, events can be described in terms of types of events (e.g., vacation, party) or holidays, places, and people. That is, one can describe an event in terms of the four W's: who, what, when, and where. The incorporated charts and columns address all but the "who" of an event. We currently provide some information about people in the form of face count per image within an event, once one or more events have been selected (see Section III.D). We plan to add other metadata based on face recognition in the future.

The month-year chart supports access to the "when" of an event, as well as quick access to annually recurring or seasonal events (e.g., recurring events typical of consumers' capture habits such as holidays, birthdays, summer vacation, etc.). For example, a user can click on the "Dec" label of the month-year chart and the event list will be filtered to show all events captured in December of any year. In addition to being able to select a month-based slice of time, users can select events from a particular year or any given month-year.

The bars in the month-year chart represent the relative number of *events* captured during any given month-year slot, and not the number of *assets* captured. However, since selection within any chart is linked to selection in all other charts, the user can simply click on the rightmost label of the event size vs. event type chart to select all events of the largest size. Making this selection would then show the distribution of all such large events over time as yellow highlighter bars within the month-year chart. Corresponding highlighted selections within the holidays chart would indicate any holidays that contain large numbers of captures. Another way to access the largest events is to click the "Show All" button above the event text list to show all top-level events and then reverse sort the list by media count. Providing multiple mechanisms for accessing or viewing details for the same type of metadata gives the researcher flexibility and context within which to analyze event data.

C. Media Display

When one or more top-level events have been selected to analyze, media from those events are displayed in the middle column (see Figure 1.). This middle column consists of:

- a horizontal slider at the top of the column for setting the relative image thumbnail size;
- a top media display area that visually shows event segmentation and hierarchy resulting from the event detection algorithm (see Section II.A);
- a horizontal gray splitter bar enabling users to change the amount of vertical space allocated to the top and bottom media display areas;
- a bottom media display area that shows individual media of the selected event(s).

Orange borders visually highlight media groupings based upon event segmentation and hierarchy in the top middle media display area of EA. The example in Figure 1. shows the event hierarchy for the six Halloween events selected. In this example, all events have event level = 1 and are thus type "Event." The orange borders indicate that the first two events contain two sub-events, the next three events contain no sub-events, and the last event contains seven sub-events.

The selection highlighting and ordering of individual media in the bottom middle area of EA are synchronized with the asset details list shown in the top right area of EA.

D. Asset-Based Metadata

Once one or more top-level events have been selected to analyze, and corresponding media are displayed in the middle column of EA, the right column is updated to display the corresponding list of assets and their metadata. Starting from the top of the right column of the example presented in Figure 1. , we see that there are a total of 53 assets (N = 53) from the potential Halloween events in the collection and that one asset is currently selected. The list view of assets presented in the asset details tab includes the asset-based metadata described in Section II.B. Interactive frequency distribution charts for face count, IVI, focal length, and child events are displayed in the bottom right of EA and summarize the distribution of assets across the corresponding metadata values.

Similar to the event details list, the asset details list can be sorted by any metadata column simply by clicking on that column's header. As mentioned in the previous section, asset thumbnails in the bottom middle area of EA are synchronized to display in the same order as the asset details list. Thus, reverse sorting the asset details list by IVI would display the highest IVI image thumbnails first and the lowest quality IVI image thumbnails last. Similarly, sorting the asset details list by face count would order thumbnails by the number of detected faces. Sorting assets by detected face count and then selecting a specific face count value (e.g., by selecting the "2" bar of the face count chart) quickly reveals the correctness of the face detection algorithm used.

Rather than filtering the list of assets to display, selections made via the asset-based charts *highlight* the corresponding assets of the asset list. Thus, clicking on any bar of any bar chart in the lower right area of EA will select the corresponding assets. Selection linking, indicated by yellow highlighting and borders, is maintained across the asset details list, the thumbnails in the bottom middle of EA, and the asset-based charts in the bottom right of EA. Thus, a user can select one or more thumbnails from the thumbnails area and the corresponding assets will be highlighted in the asset details list and yellow highlighter bars will show the corresponding metadata values in the asset-based charts. Similarly, when a user selects one or more assets from the asset details list, the corresponding thumbnails are highlighted and highlighting within the charts are updated. In the example presented in Figure 1. , the user has clicked the IVI = 2 bar and the corresponding image thumbnail containing a lower quality dark image is highlighted.

IV. Analysis Examples and Discussion

Due to limited space, it is not possible to include the full analysis of a real-life data set here, but we provide several examples in this section to illustrate how EA can be used in practice. Consider the outdoor probability chart displayed in the lower left area of Figure 1. The shape of the chart indicates that in many cases, outdoor probability at the event level can likely be used to distinguish indoor events from outdoor ones. Events with a very low outdoor probability are likely to be indoor events; events with a very high outdoor probability are likely to be outdoor events. But how correct is the outdoor probability rating? The event details list of Figure 1. shows the outdoor probability values for the selected Halloween events. The second and fourth events have the highest outdoor probability values of ~0.81 and ~0.56, respectively; the remaining values are less than 0.25

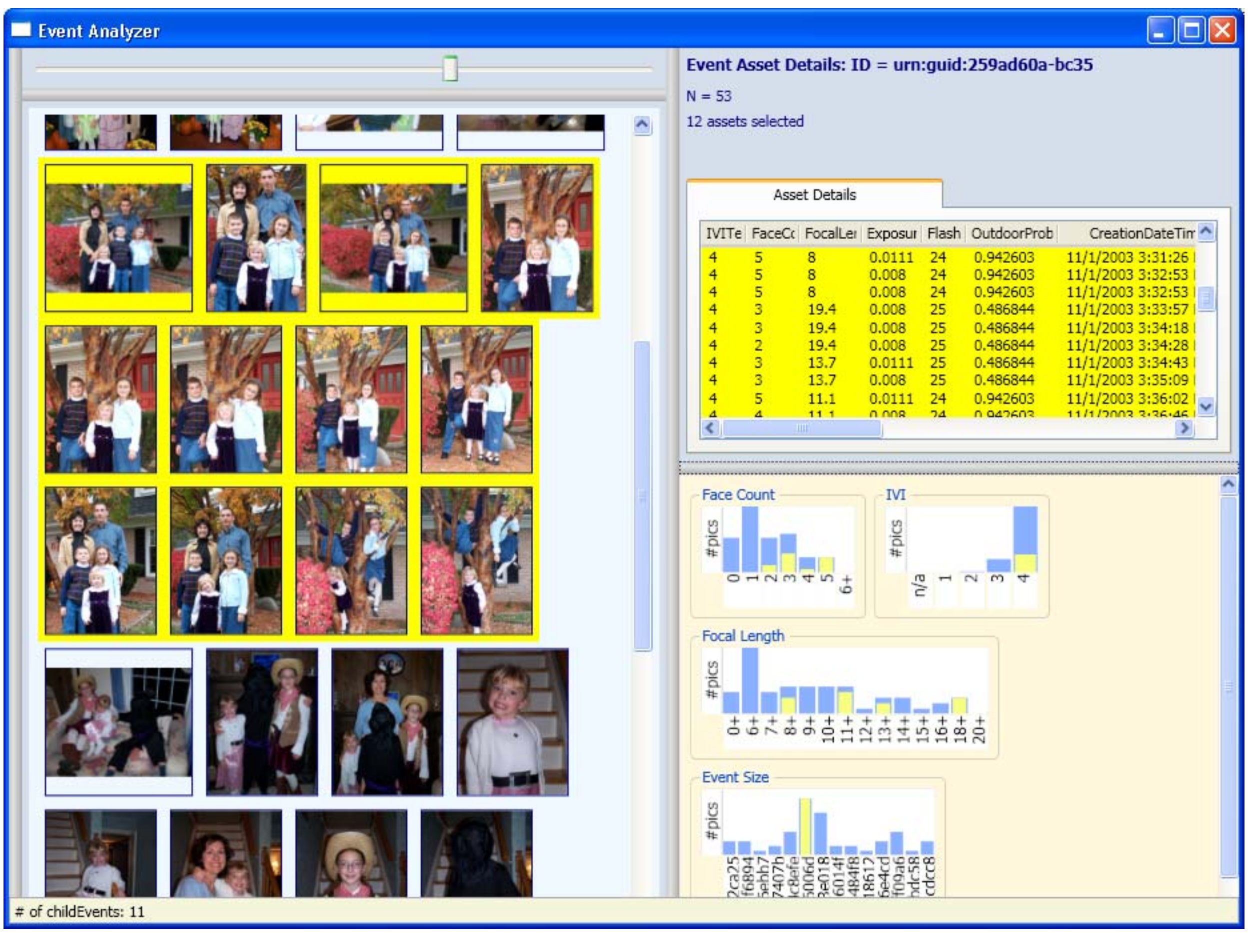

Figure 2. Sample screen shot of Event Analyzer (EA) focusing on middle and right columns.

and should thus be indoor events. The event hierarchy in the upper middle area of Figure 1. verifies that the second and fourth events do contain outdoor pictures while other events contain indoor pictures. This provides support for the utility of the outdoor probability rating of events. A researcher can select and review other events to further investigate the correctness of the outdoor probability rating of events.

Our holiday algorithm for events may add holiday tags to events that occur *near* a holiday, not just events that occur *on* that holiday specifically. The holiday probability for an event indicates the likelihood that the event occurred *on* a specific holiday. Is it useful to tag events occurring near holidays? Referring back to Figure 1. , we see that only two out of six events tagged with Halloween actually occurred on Halloween. Are the other four Halloween-tagged events related to Halloween? The event hierarchy indicates that all but one of the events, the one captured on November 1, 2003, looks Halloween-related. In this case, tagging events that occur near holidays is helpful more often than not.

We can better focus in on individual thumbnails and asset-based metadata by dragging the grid splitters (i.e., gray divider bars) to hide the left column and upper middle area (see Figure 2.). In Figure 2. , the one non-Halloween event is selected to highlight the corresponding set of outdoor family photos. Five of the highlighted photos contain all five family members; the other seven photos contain only the three children. Reviewing the face count chart indicates that of the selected photos, some contain two, three, four or five detected faces. This indicates that the face detection algorithm did not detect all faces in all photos. However, in this case, the face detector was more often correct than incorrect. In addition, the photos where one face was not detected are in the same event cluster as the photos where all faces were detected. This indicates that event clustering could be leveraged to aid users in finding related photos, once one photo within the event cluster has been found.

Researchers can continue to explore data sets in this way to gain a sense of the general utility of various metadata for any given image collection. Analyses may reveal that some algorithms perform better than others on certain types of image collections. Exposing the relationships and trends of all metadata in a common framework enables researchers to gain better understanding of the context within which algorithms perform better. This in turn could lead to new algorithm enhancements.

V. Related Work

Many researchers have proposed different algorithms for event detection and segmentation and also presented different visualizations or graphical user interfaces (GUI's) for presenting events to users (e.g., [2], [4]). While these systems provide an event-centric view of an image collection, they limit the display of event-based metadata to date- and event-size- related information and representative photo(s) for events. They also do not expose algorithmically derived asset-based metadata to the user.

Other researchers have integrated data visualization with image queries and image visualizations to aid users in interactively browsing their collections (e.g., PhotoFinder system [5]). In our work, we also combine data visualization with corresponding displays of event and image retrieval results. Our approach, however, focuses on a tool to aid researchers in visually analyzing trends and correctness of extracted and derived metadata of events and assets in an event-centric manner. In contrast, PhotoFinder was designed to aid consumers in browsing and retrieving images and appears to be heavily based upon manually input metadata [5]. Furthermore, PhotoFinder charts appear to be displayed one at a time via a tabbed interface, making it impossible to view more than one chart of metadata at once, let alone compare selections across charts.

VI. Conclusion

New tools are needed to evaluate the utility and correctness of algorithmically derived image- and event-based metadata of real-life consumer image collections. In this paper, we presented Event Analyzer, an interactive analysis tool designed to meet this need. We described how we integrated data visualization with event and image retrieval into an event-centric view of a consumer image collection for easily browsing and evaluating collection trends and metadata correctness. We also provided examples to illustrate how EA can be used in practice with real-life data. In the future, we plan to use results from EA as input towards identifying how and when imperfect algorithms can best be leveraged to aid users in browsing and searching their image collections without requiring users to manually input or verify metadata tags.

References

[1] Das, M. and A. Loui, "Event Classification in Personal Image Collections," *ICME Workshop on Media Information Analysis for Personal and Social Applications*, New York, July 2009.

[2] Graham, A., Garcia-Molina, H., Paepcke, A. and T. Winograd, "Time as Essence for Photo Browsing through Personal Digital Libraries," *JCDL'02 Proceedings*, Portland, OR, 2002, pp. 326-335.

[3] http://www.geonames.org.

[4] Huynh, D., Drucker, S., Baudisch, P., and C. Wong, "Time Quilt: Scaling Up Zoomable Photo Browsers for Large, Unstructured Photo Collections," *Ext. Abstracts CHI'05*, ACM Press, 2005, pp. 1937-1940.

[5] Hyunmo, K. and B. Shneiderman, "Visualization Methods for Personal Photo Collections: Browsing and Searching in the PhotoFinder," *IEEE International Conf. on Multimedia and Expo*, New York City, New York, 2000, pp. 1539-1542.

[6] Loui, A. and A. Savakis, "Automated Event Clustering and Quality Screening of Consumer Pictures for Digital Albuming," *IEEE Transactions on Multimedia*, vol. 5, 2003, pp. 390-402.

[7] Loui, A., Wood, M. D., Scalise, A., and J. Birkelund, "Multidimensional Image Value Assessment and Rating for Automated Albuming and Retrieval," *IEEE International Conference on Image Processing*, San Diego, October 2008.

[8] Luo, J. and M. Boutell, "Natural Scene Classification Using Overcomplete ICA," *Pattern Recognition*, vol. 38, 2005, pp. 1507-1519.

[9] Rodden, K. and K. Wood, "How Do People Manage Their Digital Photographs?" *CHI'03 Conference Proceedings*, ACM Press, 2003, pp. 409-416.

[10] Wood, M. D., "Matching Songs to Events in Image Collections," *IEEE International Conference on Semantic Computing*, Berkeley, California, September 2009.