

A Tool for Authoring Unambiguous Links from Printed Content to Digital Media

Andreas Girgensohn¹, Frank Shipman², Lynn Wilcox¹,
Qiong Liu¹, Chunyuan Liao¹, Yuichi Oneda³

¹FX Palo Alto Laboratory, Inc.
3400 Hillview Avenue
Palo Alto, CA 94304, USA

²Department of Computer Science
Texas A&M University
College Station, TX 77843, USA

³Fuji Xerox
6-1, Minatomirai, Nishi-ku
Yokohama, 220-8668, Japan

{andreasg, wilcox, liu, liao}@fxpal.com, shipman@cs.tamu.edu, oneda.yuichi@fujixerox.co.jp

ABSTRACT

Embedded Media Markers (EMMs) are nearly transparent icons printed on paper documents that link to associated digital media. By using the document content for retrieval, EMMs are less visually intrusive than barcodes and other glyphs while still providing an indication for the presence of links. An initial implementation demonstrated good overall performance but exposed difficulties in guaranteeing the creation of unambiguous EMMs. We developed an EMM authoring tool that supports the interactive authoring of EMMs via visualizations that show the user which areas on a page may cause recognition errors and automatic feedback that moves the authored EMM away from those areas. The authoring tool and the techniques it relies on have been applied to corpora with different visual characteristics to explore the generality of our approach.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems.

General Terms

Algorithms, Design, Human Factors.

Keywords

Vision-based paper interface, authoring tool, document retrieval.

1. INTRODUCTION

Embedding links to on-line content in printed documents is one approach to bridging the gap between people's consumption of digital and paper content. Such links need to be easy to use, highly accurate, and minimize the disturbance of the aesthetics of the printed medium. Our work on Embedded Media Markers (EMMs) is an approach to achieving these goals [5]. To indicate to users the presence of links and to avoid accuracy issues inherent in pure content-based retrieval approaches, EMMs are deliberately authored and are visualized on the page. To address issues of aesthetics, EMMs are semi-transparent overlays included in printed documents that make use of the existing content of the printed material

*Area Chair: Dick Bulterman

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM'11, November 28–December 1, 2011, Scottsdale, Arizona, USA.

Copyright 2011 ACM 978-1-4503-0616-4/11/11...\$10.00.

as a “fingerprint” that is used to identify the EMM during retrieval (see Figure 1). Readers photograph the printed matter including the EMM to retrieve the associated digital content.

Our initial implementation of EMMs showed much potential but showed that authors of documents embedding EMMs need support to be sure that the EMMs authored are distinct enough that they can be accurately identified. This paper describes a new EMM authoring interface that leverages the analysis of the printed content and already authored EMMs in order to provide feedback and support for authors placing EMMs in a document or set of documents.

The next section describes related and prior work. This leads into a presentation of the design of a new EMM authoring interface. We conclude with a discussion of our findings and future work.

2. RELATED WORK

There are two common approaches for creating links between paper documents and digital content. One approach is to print markers on the document. Examples are 2D bar codes [8] or printed grids of dots [3]. The second approach uses content analysis to recognize the particular portion of a paper document and determine the link based on that content. The second common approach to linking between paper and digital content is to compute features of the document content itself that identify the document patch in order to create a media link. HotPaper [1] and Mobile Retriever [6] use features based on document text such as the spatial layout of words. Other systems such as Bookmarkr [4] and MapSnapper [2] use pixel level image features, such as SIFT [7], to recognize generic document content. With these systems, visually obtrusive marks are not required for identification. As a result, in most content-based retrieval systems, there is no on-paper indication to the user that there is media linked to the document.

Embedded Media Marker indicating linked content

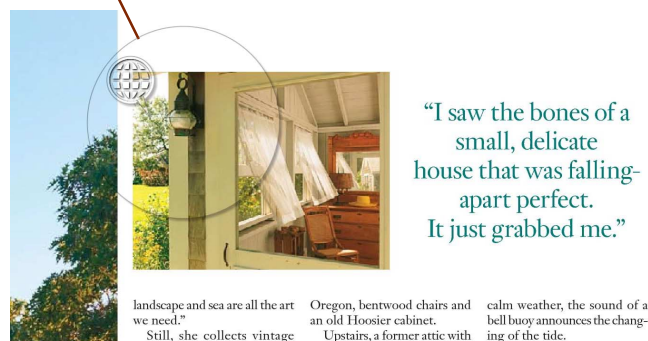


Figure 1. Part of a page with EMM as seen by the reader.

EMMs combine the characteristics of barcode-based and visual content-based approaches to improve their usability and accuracy while reducing their impact on the aesthetics of the paper document. EMMs rely on deliberately placed markers, like barcode and glyph approaches, but do not rely on encoding information in the marker (see Figure 1). Instead, the marker indicates to the reader the location and size of the link anchor. The destination of the link is determined via the content of the paper document in the region of the link. By only indexing subregions of the document page, this approach reduces the size of the index for the document. In contrast with barcodes and glyphs, the EMM icon indicates the type of media associated with the marker.

3. EMM AUTHORIZING SUPPORT

Authoring an EMM is a process of identifying a location in a collection of pages that is a source anchor for a link and the media content to which it links. Support for the authoring process is needed to help users create useful and unambiguous EMMs. We designed a web-based authoring tool to aid EMM placement and assuring that the EMM can be accurately recognized.

The authoring tool supports the user in placing EMMs on document pages to ensure recognition and to avoid conflicts. It first lets users select the page to author and then matches the keypoints on the page to those on other pages. The authoring tool uses that information to identify page regions that have a sufficient density of keypoints that would not be confused with previously authored EMMs or pages in the collection. Because of the way the recognition process works, regions that contain a group of keypoints matching the same EMM are especially undesirable. The authoring tool uses this information to guide the user with visualization and assistive techniques.

There are several components of the authoring tool. The top level manages a collection of document pages by displaying page thumbnails that a user can select for authoring. Selected pages are displayed in a center region, and can be zoomed for easier authoring. A sidebar supports creation of media links by displaying thumbnails representing links that the user can drag onto the page being authored. The authoring component allows users to place an EMM at a desired location in the page while indicating whether or not that location can be recognized accurately (see Figure 2).

The interactive authoring tool uses shading to indicate page areas that are unsuitable for EMMs. In addition to the shading, users may also turn on the display of individual keypoints that are colored in shades between blue (good) and red (bad) to indicate how well they match keypoints on other pages (see Figure 3). In the tool, a user clicks on a desired location of the page and an EMM is automatically placed near that location optimizing the recognizability. The size of the EMM is determined by a score computed from the keypoints in the region. At this point, the user can drag the EMM to other more desirable areas of the document.

An authored collection of EMMs comes from a corresponding collection of document pages. Such a collection may consist of the pages of a single book or an issue of a magazine. A collection may also encompass the pages of all the issues of a magazine published during a time period. In the latter case, the collection may change over time when new issues of the magazine are added. Readers of such books, magazines, or other documents start by identifying the



Figure 2. EMM authoring interface indicating unsuitable areas in shades of brown and a newly authored EMM in green.

collection. They then take a picture of an EMM using a portable device such as a smartphone that uploads the picture to the EMM server which redirects the device to the linked content.

The next sections describe how the authoring tool supports the authoring of unambiguous EMMs by determining suitable page regions, visualizing them, and guiding the user towards them.

3.1 Adjusting EMMs for Better Recognition

Once EMM authors select a page for placement of a new EMM or editing of an existing EMM, they are presented with a zoomable view of the page that shows the existing EMMs and an overlay indicating areas to avoid (shown in brown) when placing EMMs, as shown in Figure 2. EMMs are placed by clicking on the page image. They are visualized as by translucent green circles.

When an EMM is first placed, the smallest area with a sufficient keypoint score that contains the click position is identified. This ensures that authored EMMs includes enough distinguishable keypoints to satisfy the retrieval algorithm. Unlike the approach of simply counting the number of keypoints [5], a score is computed that measures potential confusion with other EMMs in the database. This score represents both the presence of keypoints and the confusability with other EMMs and other pages in the database.

The authoring interface is interactive, providing visual feedback of the location, size, and uniqueness of the EMM as the user moves the mouse to the desired link anchor. As the EMM is dragged, its size increases or decreases based on the score, and is shown in red when the score is below the recognition threshold and green otherwise. The computation of the score is optimized such that it can be recomputed quickly enough in JavaScript to give immediate feedback to the user in the middle of a drag operation. A lower bound of 2 by 2 inches is enforced for the region to accommodate the minimal focal distance of mobile phone cameras. If the resulting regions would be very large, the authoring tool attempts to move the EMM to avoid unsuitable regions, such as blank portions of pages. If no suitable region can be found, the EMM is colored red to indicate its unsuitability. Both assistive techniques, i.e., resizing and moving EMMs, can be turned off by the author. If the score

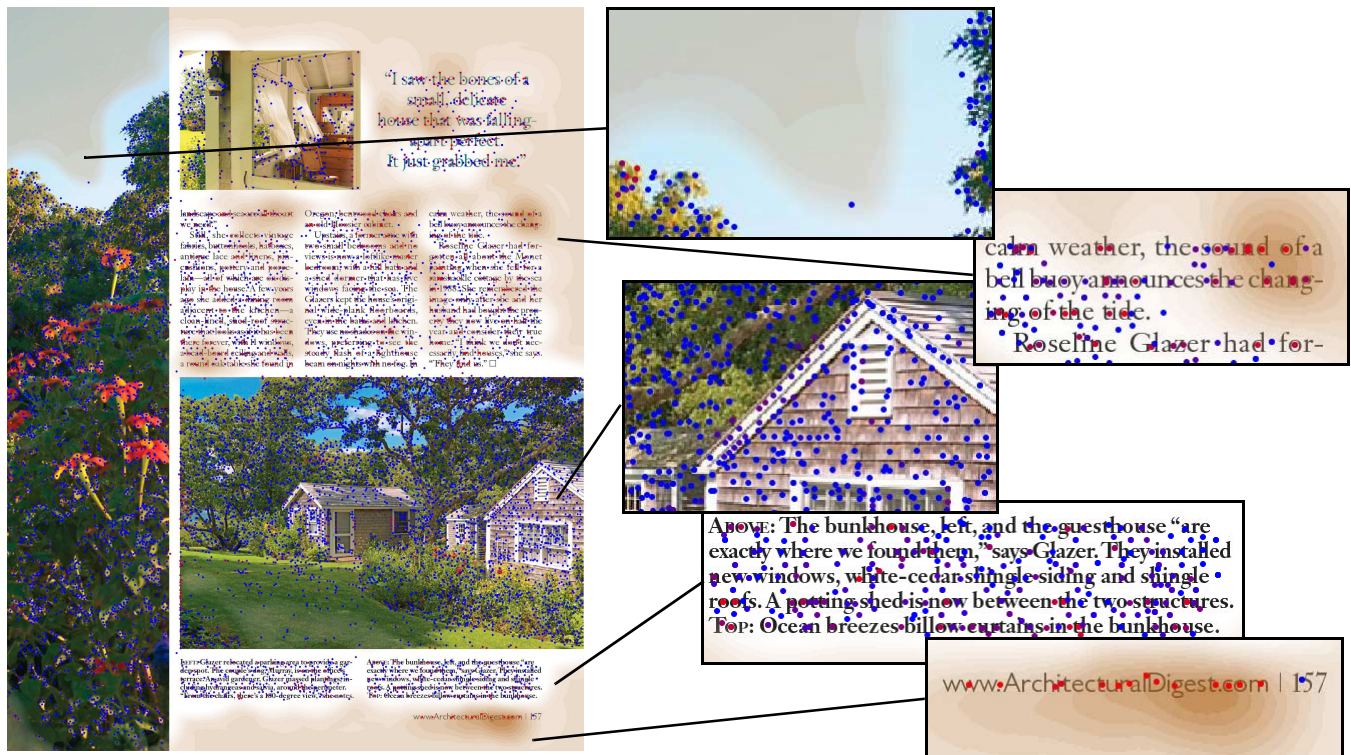


Figure 3. Visualization of keypoint conflicts and unsuitable page regions.

represents just the presence of keypoints, this support prevents the user from placing an EMM in a region with too few keypoints. However, it does not guarantee recognition accuracy. Other EMMs in the database may be associated with very similar keypoints. The following section discusses how to detect such conflicts.

3.2 Detecting Potential Conflicts

Rather than just counting keypoints in a region to determine the total score, the authoring tool determines how distinguishable those keypoints are from keypoints already in the database. The tool checks both against previously authored EMMs for actual conflicts and against other pages in the document page collection for potential conflicts. Each keypoint is given a score based on how well it matches keypoints of other EMMs and document pages. The scores of keypoints are taken into consideration when determining a suitable location for an EMM and when automatically adjusting an EMM after a user interaction. For each keypoint, the distance to the closest keypoint among keypoints in other EMMs and pages is determined. If this distance is zero, the keypoint is given a score of negative one, indicating that it does not provide good discrimination. As the distance to the neighboring keypoint increases, the score increases to a maximum of positive one. Other keypoints are assigned a score based on the distance and the maximum match distance. The normalized distance between 0 and 1 is linearly mapped to a score between -1 and 1.

When evaluating a region of a page during authoring, the scores of keypoints in that region are added. A large positive sum indicates that the region is suitable for placing an EMM. A neutral or negative sum indicates that the region is unsuitable. We use a threshold to determine if a page region for a potential EMM contains enough good keypoints. For example, we may require that the page region

contains keypoints with a total score of at least 100. By using the sum of keypoint scores instead of the keypoint count, we incorporate the notion of conflicts with other pages and EMMs. For example, either 100 keypoints with a score of 1 or 200 keypoints with an average score of 0.5 would be sufficient for a page region.

Placing two EMMs too close to each other is detrimental for retrieval because keypoints from the undesired EMM may be included in the photo of the desired EMM. While keypoint scores could be used to discourage authors from placing EMMs too close to each other on the same page by using the score density, it is simpler to check for sufficient spacing and to alert the author.

Comparing the EMM authoring page against the other pages in the collection has both advantages and disadvantages. An advantage is that regions that are ambiguous, such as text passages in the default font, can be avoided. The same words appear on many different pages so that another EMM may provide a better match to a low-quality picture of an EMM containing those words. Such regions can be found without requiring the authoring of previous EMMs. On the other hand, such an approach, if used to block the authoring of EMMs, may prevent attaching an EMM to a distinctive region that is repeated across pages.

Instead of blocking the user from authoring potentially ambiguous EMMs, the system provides support for authoring multiple source anchors for the same destination. When a user authors an EMM for a region that is repeated some number of times, the system can ask if the EMM should be added for all instances of this region. This makes it possible to map all instances of a corporate logo to the same destination and make it clear to the author that visually similar EMMs must link to the same location. The retrieval of regions similar to a selected region is a two-step process. First, all pages

with a large number of matching keypoints are located. Second, on each of those pages, regions containing many matching keypoints are located. Those regions are presented as candidates to the user.

3.3 Spreading Penalty Scores

When multiple keypoints matching the same previous or potential EMM are located in the region being considered for EMM placement, there is a much higher chance that the current location and that other location will be confused during retrieval. To reduce such conflicts, the nearby keypoints that map to the same EMM are further reduced in their score.

We call the difference between the perfect score of 1 and the actual score of a keypoint the penalty score. As the keypoint score lies between -1 and 1, that penalty score ranges from 0 to 2. The system spreads this penalty to nearby keypoints that match the same EMM. The penalty score is weighted by a distance function and the weighted sum of matching penalty scores is accumulated at each keypoint. This weighted sum of penalty scores is subtracted from the original score of the keypoint.

The spread penalty score is weighted by a normalized distance within a given radius. We use 25% of the radius of a typical EMM as the spread radius. The spread penalty score is accumulated separately and added to each keypoint score in a separate step to avoid spreading beyond the distance threshold or spreading back and forth. As shown in Figure 3, keypoint scores of -1 are visualized as red dots, scores of 1 as blue dots, and scores in between as shades of purple. Depending on the source material, other colors can be used to be noticeable on the displayed page. While keypoint dots are not shown by default, the user can turn them on to get a more detailed view of potential conflicts.

3.4 Visualizing Suitable Page Regions

In the authoring tool, the document page is displayed with an overlay that gradually shades out regions of the page with low keypoint density scores. Regions suitable for EMMs contain many keypoints not similar to those used by other EMMs and few keypoints similar to those already used. The average keypoint score density in a region indicates its suitability. A threshold for this density is used for visualization. We experimentally determined that a score threshold of 100 for a 2-inch by 2-inch region produced good results and converted that to a density threshold.

To determine the average keypoint density while taking keypoint scores into consideration, we compute a density score over a circular region with a weight that linearly decreased with distance. We use 50% of the radius of a typical EMM for the circular region. The integral under the two-dimensional curve should be normalized to one with this function.

This approach can be used with the negative keypoint scores described earlier. When combining positive and negative scores, both an empty region and a region with an equal amount of positive and negative scores would have a neutral suitability. When all keypoints have a score of 1, any region where the keypoint density exceeds the threshold has the maximum suitability. A higher density of keypoints with lower positive scores would also lead to the maximum suitability. Conversely, a high density of keypoints with negative scores produces a region with a negative suitability.

When spreading the penalty score to keypoints in the same vicinity matching the same EMM, keypoint scores can drop below -1.

When determining the keypoint density using the approach described above, such scores would cause a larger part of the circular region to be visualized as completely unsuitable, indicating to the user that this region should really be avoided. Figure 3 illustrates this visualization technique using shades of brown and different amounts of transparency. The close-ups of Figure 3 demonstrate that empty areas such as the sky are covered by a semi-translucent overlay. In contrast, the picture of the house contains a high density of blue dots representing keypoints that match nothing else so that no overlay is added. The body text at the top of the page has regions with predominately red dots that are covered with a solid overlay. This is even more the case for the page footer that only contains red dots. The figure caption with bold text does not match much else so that it would be suitable for placing an EMM.

4. CONCLUSIONS

In this paper, we motivated why unambiguous marks linking paper to digital documents are important. We described cases where Embedded Media Markers (EMMs) could be confused with one another if they were authored on top of similar content or near each other. We presented an authoring tool that uses a combination of visualization and assistive techniques to guide the user. It steers users away from potentially confusing page regions. It does this by analyzing previously authored EMMs and other pages in the document collection to identify regions in the page being authored that can be confused with keypoints in the database.

EMMs already provide a means for placing links on paper that offer excellent retrieval performance while indicating the presence of links with minimal impact on the document aesthetics. We expect that the support for simple authoring of unambiguous EMMs will make EMMs even better suited for linking from paper documents to associated digital content.

5. REFERENCES

- [1] B. Erol, E. Antunez, J. Hull. HOTPAPER: multimedia interaction with paper using mobile phones. *Proc. of Multimedia '08*, 399-408, 2008.
- [2] J. Hare, P. Lewis, L. Gordon, G. Hart. MapSnapper: Engineering an efficient algorithm for matching images of maps from mobile phones. *Proc. of Multimedia Content Access: Algorithms and Systems II*, 2008.
- [3] B. Hecht, M. Rohs, J. Schöning, A. Krüger. Wikeye – Using Magic Lenses to Explore Georeferenced Wikipedia Content. *Proc. of Pervasive Mobile Interaction Devices Wrkshp*, 2007.
- [4] N. Henze, S. Boll. Snap and share your photobooks. *Proc. of Multimedia '08*, 409-418, 2008.
- [5] Q. Liu, C. Liao, L. Wilcox, A. Dunnigan, B. Liew. Embedded media markers: marks on paper that signify associated media. *Proc. of Intelligent User Interfaces*, 149-158, 2010.
- [6] X. Liu, D. Doermann, Mobile Retriever: access to digital documents from their physical source. *Int. Jour. on Document Analysis and Recognition*, 11(1): 19-27, 2008.
- [7] D.G. Lowe. Distinctive image features from scale-invariant keypoints, *Int. Jour. of Computer Vision*, 60(2), 91-110, 2004.
- [8] J. Rekimoto, Y. Ayatsuka. CyberCode: designing augmented reality environments with visual tags. *Proc. of DARE 2000*, 1-10, 2000.