Demonstrating Reality-Based Information Retrieval

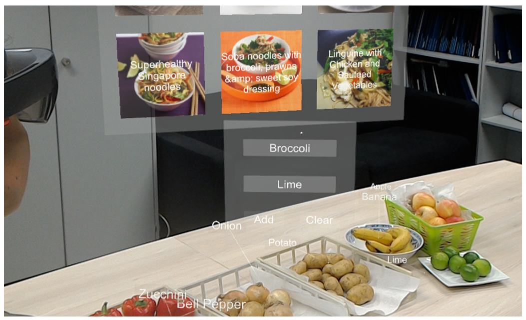

Figure 1: Recipe search with our RBIR prototype. Fruits and vegetables are labeled and serve as input for the query formulation. Matching recipes are displayed. Result data retrieved from Food2Fork.com

Wolfgang Büschel
Interactive Media Lab
Technische Universität Dresden
Dresden, Germany
bueschel@acm.org

Annett Mitschick
Interactive Media Lab
Technische Universität Dresden
Dresden, Germany
annett.mitschick@tu-dresden.de

Raimund Dachselt
Interactive Media Lab
Technische Universität Dresden
Dresden, Germany
dachselt@acm.org

CHI'18 Extended Abstracts, April 21–26, 2018, Montreal, QC, Canada
ACM 978-1-4503-5621-3/18/04.
https://doi.org/10.1145/3170427.3186493

Abstract
With this work, we demonstrate our concept of Reality-Based Information Retrieval. Our principal idea is to bring Information Retrieval closer to the real world, for a new class of future, immersive IR interfaces. Technological advances in computer vision and machine learning will allow mobile Information Retrieval to make even better use of the people's surroundings and their ability to interact with the physical world. Reality-Based Information Retrieval augments the classic Information Retrieval process with context-dependent search cues and situated query and result visualizations using Augmented Reality technologies. We briefly describe our concept as an extension of the Information Retrieval pipeline and present two prototype implementations that showcase the potential of Reality-Based Information Retrieval.

Author Keywords
Reality-Based Information Retrieval; Augmented Reality; Spatial User Interface; Immersive Visualization; In Situ Visual Analytics

ACM Classification Keywords
H.5.1 [Information interfaces and presentation: Multimedia Information Systems]: Artificial, Augmented, and Virtual Realities

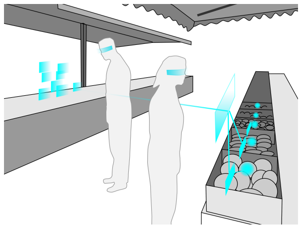

Figure 2: Illustrative example for RBIR: Two users explore a farmer's market. Information about the items is displayed, matching recipes are suggested.

Figure 3: Our first prototype: Here, a yellow jug is photographed (center), surrounded by the detected tags/properties. Results retrieved from Pixabay.com.

Figure 4: By combining two tags, the user formulated a search query for pictures matching "yellow window".

Introduction & Background

Information Retrieval plays a dominant role in our daily lives and we mostly rely on search engines to fulfill our information needs. Today, with the dissemination of powerful mobile devices and appropriate bandwidth, search has gone mobile: Statistics show that today more web searches take place on mobile devices than on computers[1]. We believe that in the future, Information Retrieval (IR) will not only be mobile and location-based but tightly interwoven with the *physical world* itself in what we call *Reality-Based Information Retrieval (RBIR)* [3].

We aim to extend and support a broad range of search applications with the help of Augmented Reality (AR) technology by providing Natural User Interfaces based on the situated visualization of search stimuli, queries, and results. Search facilities will be a vital part of future mixed reality glasses, not only because of the constant need for information in mobile contexts but also because of their novel opportunities. For example, they allow easy Just-in-time Information Retrieval (JITIR) [8] and support serendipity in information access by fusing virtual and physical information artifacts and suggesting contextual information [1].

As of yet, little work has been done to address issues in the intersection of the fields of IR and AR. Examples include [1], a prototype system which retrieves information relevant to contextual cues, i. e., people's IDs based on face recognition. Similarly, in [7] face recognition techniques were used to retrieve corresponding video snippets from a personal lifelog. General visual cues as input parameters, showing the feasibility of content-based query term extraction from photos, where shown, e. g., in [6]. Also, concepts such as *In Situ Visual Analytics* [5] or *Situated Analytics* [4] are related to our concept of RBIR. A detailed review of the related work can be found in our paper [3].

[1]Google Blog, May 2015: https://adwords.googleblog.com/2015/05/building-for-nextmoment.html

Basics of Reality-Based Information Retrieval

The basic idea of RBIR is to extend the Information Retrieval pipeline by including the three elements of an AR application described by Billinghurst et al. [2]: real physical objects, virtual elements, and interaction metaphor. These three components are represented in RBIR as the *physical world and its stimuli*, *situated augmentations*, and *natural interaction* (Figure 5) and form a new class of IR interface.

Both specific real-world objects, their properties, and the environment in general are one of the main triggers for information needs, as illustrated in Figure 2. In RBIR, these real-world object properties are represented as *Situated Stimuli*. They include low-level features such as color and textures but also higher-level concepts like object classes or even data and services associated to an object or location. Ideally, Situated Stimuli correspond directly to the psychological stimuli from the physical world that trigger an information need and serve as input for a search query.

Using natural interaction techniques, e. g. gestures, the queries are specified and visualized as *Situated Query Representations*. In the same way, *Situated Result Representations* are embedded into the physical world to visualize the results. They can then be explored by the users, which also allows for, e. g., gesture-based relevance feedback or digital annotations using voice input.

For more details on the concept of RBIR, our initial evaluations, and open research challenges, we refer to our paper [3] and the project website[2].

[2]RBIR project website: https://imld.de/rbir/

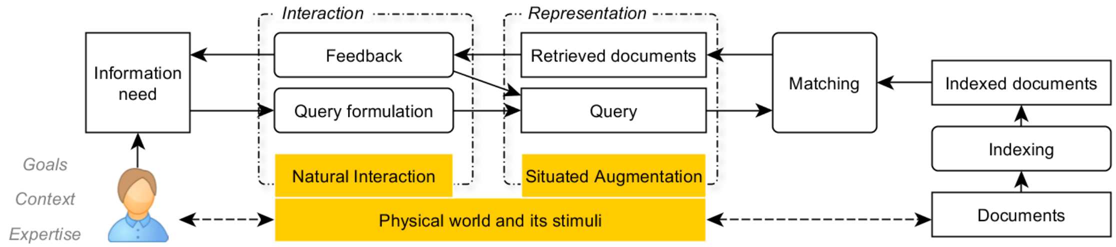

Figure 5: Conceptual model of Reality-Based Information Retrieval. A typical Information Retrieval process is extended by adding the user and three aspects of AR: the physical world, situated augmentations, and natural interaction.

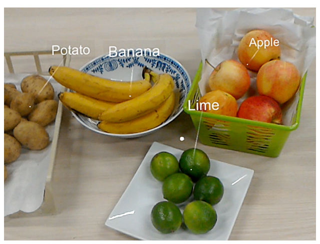

Figure 6: Labeled fruits and vegetables in the second prototype (recipe search).

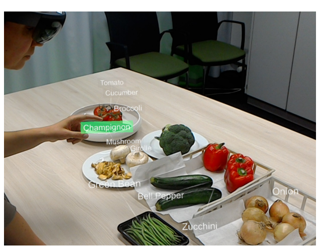

Figure 7: User selecting tags in the recipe search prototype.

Prototypes

In this demonstration we present two prototypes that show the general feasibility of our concept and implement the three main aspects of RBIR: *natural interaction techniques* to formulate a query based on physical, *real-world stimuli*, and *situated augmentations* enriching the user's environment. They both realize different information retrieval concepts for AR glasses such as the Microsoft HoloLens, using its built-in functions for image processing, augmentation, and interaction.

Situated Photograph Image Retrieval

Our first prototype implements the basic concept of situated photograph queries in AR for image retrieval. It uses a photograph metaphor to extract visual information from the environment as query parameters. Performing an air-tap gesture, the user takes a photograph of the scene (Figure 3). This photo is displayed as a query object in the middle of a 2D canvas located in front of the user.

It is sent to the Microsoft Computer Vision API to retrieve automatically assigned tags. Based on these tags, preview images are retrieved from the Pixabay.com API. The retrieved images and their labels are displayed around the initial query image. The user can tap on a tag to search for images with the specific tag. A second 2D canvas, freely placed on an available real-world surface, serves as result visualization. This result canvas shows the sorted set of images retrieved from the Pixabay.com API using the selected tags (Figure 4). The user can also take multiple photos and thus create multiples query canvases to combine tags corresponding to different objects. Selecting one of the preview images replaces the original photo with the chosen image, which is then used as query source, allowing the user to iteratively browse through pictures.

Recipe Search

In the second prototype, fruits and vegetables are augmented with labels (Figure 6). The user forms a search query by tapping on them (Figure 7). As results, recipes containing the selected ingredients are presented (Figure 1). Conceptually, instead of explicitly taking photographs, the current camera image is automatically analyzed in the background, showing tags as they are generated and thus better supporting serendipity. For improved reliability and stable results necessary for the demonstration, we decided to pre-generate the tags for this prototype instead of using an actual computer vision API. The recipe search is implemented using the Food2Fork.com API, which returns a list of recipes with preview images and other metadata.

The visualization for this prototype is graph oriented, with tags being connected to the corresponding real-world location and also to the query that they are a part of. This helps to show the spatial relations between keywords and results. The query and the results are shown on spatially positioned 2D canvases that automatically turn towards the user. Similar to the first prototype, interaction is based on gaze pointing and the air tap gesture of the HoloLens. Users can select tags to add them to the query or simply drag them to the canvas.

Conclusion

In this demonstration we present our prototypes for the concept of Reality-Based Information Retrieval. Information needs are often stimulated by the user's surroundings. Thus, we combined the Information Retrieval process with Augmented Reality by extending the IR model with the notions of the physical world, natural interaction, and situated augmentations. We showcase two example implementations for AR image retrieval and recipe search that show the feasibility of our concepts and provide a glimpse into a possible future for Information Retrieval.

Acknowledgements

We would like to thank Philip Manja for his contributions, including his work on the first prototype. This work was funded in part by grant no. 03ZZ0514C of the German Federal Ministry of Education and Research (measure Twenty20 – Partnership for Innovation, project Fast).

REFERENCES

1. Antti Ajanki, Mark Billinghurst, Toni Järvenpää, Melih Kandemir, Samuel Kaski, Markus Koskela, Mikko Kurimo, Jorma Laaksonen, Kai Puolamäki, Teemu Ruokolainen, and others. 2010. Contextual information access with augmented reality. In *Proc. 2010 IEEE Int. Workshop MLSP*. IEEE, 95–100.
2. Mark Billinghurst, Adrian Clark, and Gun Lee. 2015. A Survey of Augmented Reality. *Found. Trends Hum.-Comput. Interact.* 8, 2-3 (March 2015), 73–272. `http://dx.doi.org/10.1561/1100000049`
3. Wolfgang Büschel, Annett Mitschick, and Raimund Dachselt. 2018. Here and Now: Reality-Based Information Retrieval. In *Proc. CHIIR '18*. ACM, New York, NY, USA, to appear. `http://dx.doi.org/10.1145/3176349.3176384`
4. Neven ElSayed, Bruce Thomas, Kim Marriott, Julia Piantadosi, and Ross Smith. 2015. Situated Analytics. In *2015 Big Data Visual Analytics (BDVA)*. 1–8. `http://dx.doi.org/10.1109/BDVA.2015.7314302`
5. Barrett Ens and Pourang Irani. 2017. Spatial Analytic Interfaces: Spatial User Interfaces for In Situ Visual Analytics. *IEEE Comp. Graph. Appl.* 37, 2 (2017), 66–79. `http://dx.doi.org/10.1109/MCG.2016.38`
6. Xin Fan, Xing Xie, Zhiwei Li, Mingjing Li, and Wei-Ying Ma. 2005. Photo-to-search: Using Multimodal Queries to Search the Web from Mobile Devices. In *Proc. MIR '05*. ACM, New York, NY, USA, 143–150. `http://dx.doi.org/10.1145/1101826.1101851`
7. Masakazu Iwamura, Kai Kunze, Yuya Kato, Yuzuko Utsumi, and Koichi Kise. 2014. Haven't We Met Before?: A Realistic Memory Assistance System to Remind You of the Person in Front of You. In *Proc. AH '14*. ACM, New York, NY, USA, Article 32, 4 pages. `http://dx.doi.org/10.1145/2582051.2582083`
8. Bradley J Rhodes and Pattie Maes. 2000. Just-in-time Information Retrieval Agents. *IBM Syst. J.* 39, 3-4 (July 2000), 685–704. `http://dx.doi.org/10.1147/sj.393.0685`