Layout Recognition for Dynamic Contents Retrieval in Digital Signage

Adiyan Mujibiya*

Abstract. We report our approach to bridge dynamic content transfer from publicly available digital signage to users' private smart devices, such as glass-like wearables and smartphones. We aim to address issues concerning dynamic multimedia signage that has its contents divided into several sections. This type of signage has become increasingly popular due to optimal content exposures. In contrast to prior research, our approach excludes computer vision (CV) based object recognition, and instead took a novel approach to identify how contents are being laid-out in a digital signage. We incorporate CV techniques to recognize basic layout features such as corners and line segments, which are obtained from camera frame of user's device. Consequently, the visual features are combined to generate *content layout map*, which will be compared to pre-learned layout map to aggregate corrected perspective. To grab a specific content, users are able to choose a section within the captured layout using the device's interface, which in turn creates a request to *contents server* to send respective content info based on a timestamp and a unique section ID. We describe implementation details, report preliminary results, and conclude with discussion of our experiences in implementation as well as highlighting our future plans.

1 Introduction

In recent years digital signage has penetrated significant numbers of public spaces and gradually replacing traditional printed or electric (bulbs/LED) billboards. This trend is mainly driven by the decreasing cost to deploy large LCD displays that are practically suitable to visualize dynamically changing and varied information. The common practice to maximize contents exposure in digital signage is to divide the screen canvas into several sections respective to content's type or genre. These multi-section signage offers clean layout design and support simultaneous contents delivery.

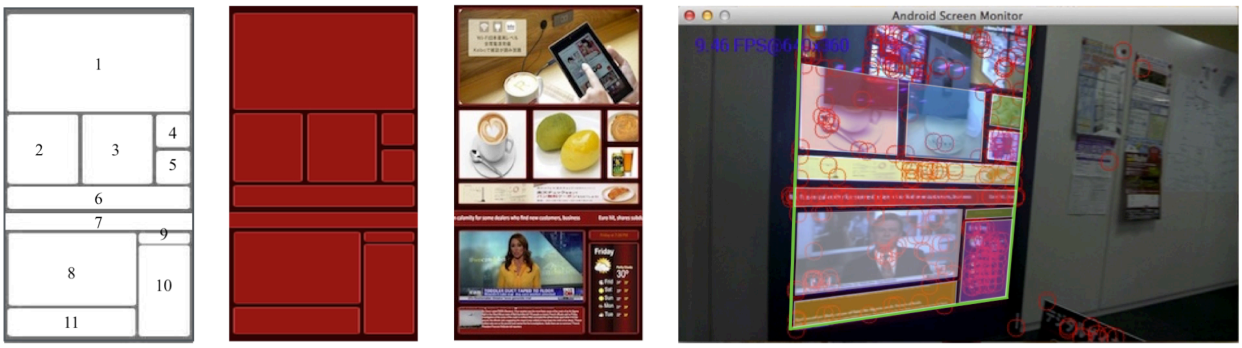
Research aiming to close the gap between public signage and user's private mobile devices has been conducted very actively in HCI field [1,2,4]. However, to the best of our knowledge, there is no work that specifically addresses contents bridging for multi-section digital signage. QR codes as well as other 2D/3D barcodes are still widely used to provide users with entry access point to detailed information. However, visual codes and URLs are incompatible for multi-section signage due to spatial and design constraints.

In this abstract, we report our approach to address aforementioned issues by implementing signage layout recognition based on CV techniques, which includes corner and line segment detection [3] of images captured on the user's device. Corners and line segments are combined with other 2D features to generate *content layout map*, which is then matched with pre-learned layout map to aggregate corrected perspective. To grab a specific content, users are able to choose a section within the captured layout using the device's interface, which in turn creates a request to *contents server* to send respective info based on a timestamp and a unique section ID. Actual usage scenario is depicted in Figure 1.

Figure 1. We report our approach to realize contents transfer from multi-section digital signage to users' private device, such as glass-like system and smartphones. We use computer vision approach, which includes corner and line segment detection to identify signage layout within user's camera view.

Using this approach, we mitigate requirement to pre-learn contents for section-specific object recognition, thus allowing dynamic changes of contents in each section, which in real-world practices may include seasonal scheduling, real-time updates, changing contents source, and so on. Moreover, our layout

*Rakuten Institute of Technology, Rakuten, Inc.
楽天株式会社 楽天技術研究所



contents layout | signage background | contents filled | viewfinder from glass-like wearable device

Figure 2. We describe our signage layout recognition system. First, we define *contents layout* to be trained using its corners and line segments that serves as a groundtruth of our layout matching. Second, we fill the contents that are positioned according to the pre-defined static layout. Here, the contents in each section can be changed dynamically. Lastly, we perform layout recognition using a camera on user's device, perform perspective correction, and let the user select a section which s/he wants to receive further information from.

recognition performs in real-time at up to 8 fps for VGA resolution (640×360), deployed in a glass-like wearable device with OMAP4430 2GB RAM dual-core SoC, on Android 4.4.2 operating system.

2 Proposed Method

We highlight our proposed method in Figure 2. We define a static layout and assign a unique ID for each section. We aggregate features such as corners and relative positioning of corners, and line segments. We use these features to form content layout map, which is being used to match features aggregated from camera frames, and obtain corrected layout perspective.

Based on the static layout, we fill the contents using standard web technologies. In our prototype, we incorporate 11 sections in our layout, and assign contents such as café poster slideshow, food and beverage menu slideshow, special offer slideshow, news update text feed and video-cast, as well as weather cast.

We implement client-server protocol to realize layout recognition and content distribution. Our layout recognition software was implemented mostly in Java, with some parts in C++ for computationally resource consuming functions. We utilize glass-like wearable as target device, considering that the device is equipped with a see-through optical head-mount display that will emphasize user experience on *public-to-private* contents retrieval. An exemplary figure of user's view is depicted in Figure 2 (right). The green outline visualizes perspective corrected signage layout, and sections are highlighted in color shades. To select a section, user can browse through sections using the glass device's touch interface. Finding appropriate interface design to accommodate section selection for specific device remains future work. After user selection has been confirmed, the client software creates a request to the content distribution server to send back a URL that refers to details or further information related to selected section. The server side of our system also manages contents scheduling, therefore pairing between timestamps to layout section IDs and its contents are straightforward.

3 Discussion and Future Work

We define content layout map based on line segments, which are locally straight zones of the image where the grey level is changing significantly. Therefore, our approach requires visible frame design (observable in Figure 2, depicted as black pixels in signage background). This potentially become a visual design constraint that limits signage designers when creating aesthetical elements in their contents. Exploring layout recognition without aforementioned constraint remains future work.

Reference

[1] She, J., Crowcroft, J., Fu, H., and Li, F.. Convergence of interactive displays with Smart mobile devices for effective advertising: A survey. *ACM Trans. Multimedia Comput. Commun. Appl.10*, 2 (2014), Article 17.a.

[2] Turner, J. Cross-device eye-based interaction. In *UIST'13*, 37-40.

[3] Von Gioi, R.G., Jakubowicz, J., Morel, J.M., Randall, G. LSD: a Line Segment Detector. In *Image Processing On Line*, 2 (2012), pp. 35–55.

[4] Yamaguchi, T., Fukushima, H., Tatsuzawa, S., Nonaka, M., Takashima, K., and Kitamura, Y.. SWINGNAGE: Gesture-based mobile interactions on distant public displays. In *ITS'13*, 329-332.