# Narrative Image Composition Using Objective and Subjective Tagging

Anusha Withana[†][*], Rika Matsui[†], Maki Sugimoto[†], Kentaro Harada[‡], Masa Inakage[†]
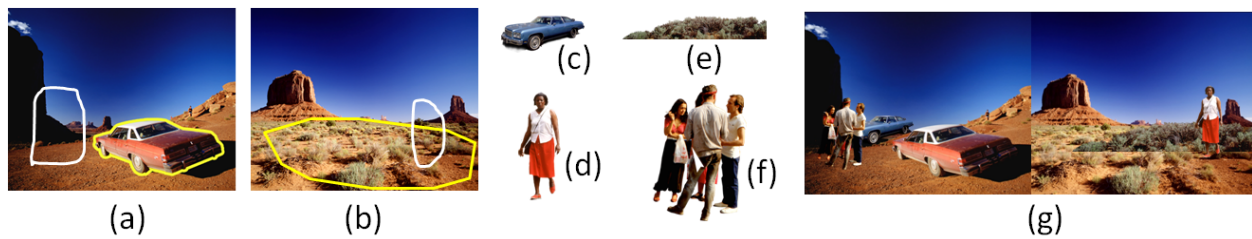† Keio University, Japan             ‡Olympus Corporation

**Figure 1:** *Shows tag information and final composite image according to tags. (a) image description: desert mountain; objective tag (yellow): car; subjective tag (white): people; (b) image description: African desert; objective tag (yellow): grass; subjective tag (white): people; (c) image description: car; (d) image description: African American women; (e) image description: grass; (f) image description: group of people; (g) Part of the final composite image, matched and composite into a single image according to the meta-data*

## 1 Introduction

With recent advancements in digital photography, data storage and network technologies, publishing and sharing of digital images in Internet has been drastically increased. High popularity and growth of internet image libraries such as Flickr and Picasa are good examples for these trends. In order to enable easy browsing and searching, online storages store meta-data in the form of keywords to describe images. Meta-data could be a general description of the image, a specific tag or an annotation to describe spatial information within the image. In order to ease and improve the efficiency of tagging process, image processing and analysis algorithms has been combined to manual tagging systems [Yang et al. 2009]. Furthermore, meta-data can be used to create interesting presentations of images. Specially in exhibition displays, automated slide shows and digital photo frames, meta data are used to group related images and present them under different categories. However, presentations generated by most of the existing systems are limited to sequential displaying of individual images as correlated groups. In this paper we present a system that composite digital images in a narrative fashion utilizing objective and subjective tagging information. Proposed system can extract a region from one image and composite it into another image according to available meta-data. As a proof of concept we created a right to left continuous image panning application using Adobe Flash which combines different images according to their similarities and composite them together to create a narrative image presentation.

## 2 Our Approach

Most of the existing image annotation systems are limited to objective tagging. For an example, users will tag or annotate contents of the image such as a visible object or an animal which exists in the target image. These information are very useful to extract information about the image, its contents and correlate them with other images with similar contents. However, they provide little help to composite an external image object between the existing content of the image. Our goal is to extract image features such as humans, animals or objects from one image and composite them into a single narrative image sequence. In order to achieve this goal, we use subjective tagging of images.

In subjective tagging, users are allowed to tag contents on the images which are probable to exist but currently not available. For an example, let us say an image has a scenery of sea and a user can mark a polygon region on the sea and tag it as a boat even though it does not actually exist on the image. It is similar to a container in the image which is suitable to hold a boat.

Subjective and objective tags are separately stored in a relational database. System uses a keyword count based Pearson correlation analysis algorithm to generate correlation indexes between each keywords and tags [Segaran 2007]. Two types of correlations are created between tagged contents. First is relationship between objective tags and subjective tags using keyword counts. Secondly, intra-image spatial correlation index to derive content types which are always closely located with each other in an image. For an example, a tag to describe sea can be highly correlated to a tag describing a ship because these two tags can coexist in multiple images.

In our image panning application, initial image to display is pre-selected from a repository. Relationship index between subjective and objective tags are used to find a possible content to fill in a region marked as a subjective tag on the selected image. Once a probable matches are found, feature marked in the objective tag is extracted by cropping out rest of the image and composite it on to the target image. Furthermore, intra-image spatial correlation index is used to composite possible other contents that could coexist in the same target image. Once the intra-image composition is finished, system calculates a matching image to attached to the right of the target image to make a continuous sequential panning of images. If possible, objective tags are used to composite a bridging feature between adjacent images. This process continues till system finishes displaying all available images.

## References

SEGARAN, T. 2007. *Programming Collective Intelligence.* O'Reilly Media.

YANG, K., WANG, M., AND ZHANG, H.-J. 2009. Active tagging for image indexing. In *Proceedings of the IEEE International Conference on Multimedia and Expo, ICME*, IEEE, 1620 – 1623.

---

[*]e-mail:anusha@kmd.keio.ac.jp