Colorific: A mixed initiative model for choosing the right color.

Chinmay Kulkarni
Stanford University HCI Group
Computer Science Department
Stanford, CA 94305
chinmay@cs.stanford.edu

Julie Fortuna
Stanford University HCI Group
Computer Science Department
Stanford, CA 94305
jfortuna@stanford.edu

ABSTRACT
Is it possible to automatically create a color palette relevant to a topic? Could such a palette be used to guide color choices while visualizing data? We envision a tool that automatically creates aesthetically pleasing and topic-relevant palettes for a large class of topics. In order to do this, we must first extract palettes from color pixel values of images from Google Images via clustering and topic models.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces.

General terms: Design, Human Factors, Experimentation

Keywords: Information visualization, colors, crowdsourcing, user study

INTRODUCTION
TODO Julie/Chinmay. 300 words. Example based– Stanford Cardinal? Anger? Sadness? Understanding is aided, if supported

RELATED WORK
CEK 400 words. One para each: Color, mixed initiative, emotions and color (http://socrates.berkeley.edu/ plab/).

Prior work exists on automatic creation of color palettes. This work falls broadly in two categories. The first focuses on finding representative colors from images, that can be used as color palettes. The most recent of these is [8]. This line of research has so far focused only on extracting colors from a single image. This project extends this work by extracting colors from multiple, related images. I believe that some of the techniques used by [8], such as a weighted histogram that uses color saturation and neighborhood color coherence, can be adapted for multiple images too. Depending on constraints of time, I plan to explore some of these techniques.

The second category of research on palette generation focuses on optimizing visual properties, such as color saliency and perceptive color distance, both manual or rule-based, as pioneered by Brewer [3]; and with varying degrees of automation [6, 9]. I believe most such optimization research is complementary to this project, and can be used as a post-extraction step to optimize the colors chosen. Statistical work on color saliency is valuable, even if it hasn't been directly applied as a optimization objective; color saliency in the context outside data-visualization in [4, 1].

Topic models have been shown to be effective in information retrieval. Latent semantic analysis (and later, LDA), for instance, has been used to find "latent" similarities between concepts [5, 2]. Similar similarity-measures have been computed for nodes in a graph [7]. While these similarity measures may help to better cluster color-values, they don't target the domain of color recommendations directly.

SYSTEM DESCRIPTION

550 Words

Images as a source of color data

Colorific obtains colors for a topic from images that are labeled to be related to the topic. The assumption here is that images that are related to a topic will contain the topic's characteristic colors. Therefore, the first step in the Colorific pipeline is to obtain a set of images that relate to a topic.

In order to do this, Colorific uses a labeled corpus that contains images along with "tags" or topics that it is related to. Several labeled corpora exist on the Internet: for example, Flickr contains primarily photographs that have been tagged manually. ImageNet contains a taxonomy of images. Google Images, and other search engines, while not an explicit tagged corpus, can also provide images relevant to a topic (through search). Colorific uses Google Images as its image source, because of the large number of images it indexes (unlike, Flickr, which consists primarily of photographs), and because it does not images to be tagged explicitly (unlike ImageNet). This increased diversity and quantity in the corpus comes at a price, however– images vary in quality, size and relevance to topic. However, Colorific is largely robust to these these problems, as described below.