

Semantic Awareness for Automatic Image Interpretation

Albrecht Lindner

Qualcomm, San Diego, USA

Advisor: Professor Sabine Süsstrunk

Date and location of PhD thesis defense: 1 March 2013, Ecole Polytechnique Fédérale de Lausanne

Received 26th Jan 2014; accepted 25th May 2014

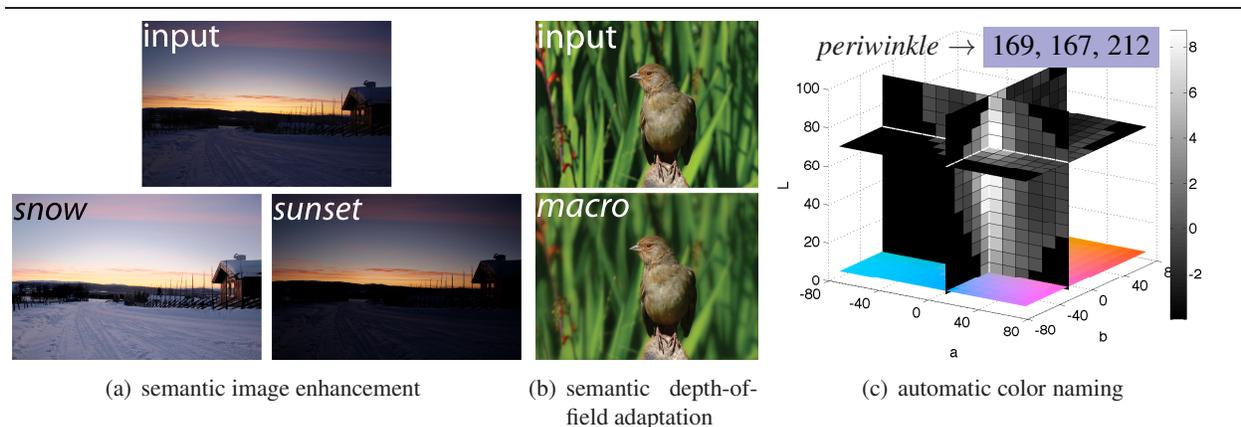


Figure 1: Example applications of the large scale statistical framework developed in this thesis. Using millions of freely annotated images from the world wide web we can (a) semantically enhance an image for a given semantic context such as *snow* or *sunset*, (b) semantically adapt an image's depth-of-field if the given keyword, e.g. *macro*, indicates a user's desire for this artistic effect, or (c) automatically determine the color values associated with a color name, e.g. *periwinkle*; the three orthogonal planes are cross-sections of a significance distribution that relates CIELAB color values to the given semantic expression.

1 Abstract

Finding relations between image semantics and image characteristics is a problem of long standing in computer vision, image analysis or related fields. Classic research in these fields is intended for applications that go from the image domain to the semantic domain such as face recognition or scene understanding. This thesis [1] explores methods and applications that go the opposite direction, i.e. use existing semantic information to infer knowledge and actions in the image domain as shown in Figure 1.

Semantic information related to images is increasingly available due to the wide spread of social online communities such as Flickr and Facebook and comes in varying forms such as tags or comments. Such semantic information is challenging because users are not restricted to a limited vocabulary but can use all words and expressions they like. Additionally, many images are necessary in order to have enough example images per keyword for significant results. Algorithms for this type of data thus have to scale very well to both large vocabularies and to large image databases.

Correspondence to: <ajl.epfl@gmail.com>

Recommended for acceptance by <Alicia Fornés and Volkmar Frinken>

ELCVIA ISSN:1577-5097

Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

The underlying mathematical foundation for this thesis is a statistical framework to relate image keywords to image characteristics and is based on a large database of annotated images. The design of the framework respects two equally important properties. First, the output of the framework, i.e. a relatedness measure, is compact and easy-to-use for subsequent applications. We achieve this by using the simple, yet effective Mann-Whitney-Wilcoxon significance test. It measures a given keyword's impact on a given image characteristic (see Fig. 1(c)), which results in significance values that serve as input for successive applications. Second, the framework is of very low complexity in order to scale to millions of images and thousands of keywords.

The first application we present is semantic image enhancement [2]. The enhancement framework takes two independent inputs, which are an image and a keyword, i.e. a semantic concept. The algorithm then re-renders the image to match the semantic concept. We implement this framework for color enhancement and depth-of-field adaptation of images. Figure 1(a) shows an example image that is enhanced for two different semantic concepts. The algorithm brightens the image in the context of *snow* as these images tend to have more white pixels, but it darkens the image in the context of *sunset* and makes the reddish colors more salient. Figure 1(b) depicts an image where the user can strengthen a desired out-of-focus blur by adding the keyword *macro*. The algorithm automatically infers this artistic intention from the keyword and accordingly increases the blur in the background while maintaining the in-focus object untouched. More example images are available on the related research webpage: <http://ivrg.epfl.ch/SemanticEnhancement.html>

Unlike conventional image enhancement algorithms, our proposed approach is able to re-render a single input image for different semantic concepts, producing different image versions at the output to reflect the image context as shown in Figure 1(a). This redefines the notion of a well-enhanced image, which is usually seen as a measure of only the image itself, whereas in our setup the optimal image characteristics depend also on the semantic context.

We evaluate the proposed semantic image enhancement with two psychophysical experiments. The first experiment comprises almost 30'000 image comparisons of the original and the enhanced images and was crowd-sourced on Amazon Mechanical Turk. The majority of the enhanced images was proven to be significantly better than the original images. The second experiment contains images that were enhanced for two different keywords. We compare our proposed algorithm against histogram equalization, Photoshop auto-contrast and the original. Our proposed method outperforms the others by a factor of at least 2.5.

The second application we present is color naming (see Fig. 1(c)), which aims at relating color values to color names and vice versa [3]. While conventional color naming depends on psychophysical experiments, we are able to solve this task fully automatically using the significance values from the statistical framework. We first demonstrate the usefulness of our approach with an example of 50 color names and then extend it to the estimation of memory colors and color values for arbitrary semantic expressions.

In a subsequent experiment, we use a list of over 900 English color names and translate them to 9 other European and Asian languages [4]. We estimate color values for these over 9000 color names and analyze the results from a language and color science point of view. The color estimations are publicly available in the form of an online color thesaurus: <http://colorthesaurus.epfl.ch>. Users can explore color names by browsing through color space and different languages.

References

- [1] A. Lindner, *Semantic Awareness for Automatic Image Interpretation*, PhD thesis, EPFL, 2013.
- [2] A. Lindner, A. Shaji, N. Bonnier and S. Süsstrunk, "Joint Statistical Analysis of Images and Keywords with Applications in Semantic Image Enhancement," *ACM Multimedia*, Osaka, 489 – 498, 2012.
- [3] A. Lindner, N. Bonnier and S. Süsstrunk, "What is the color of chocolate? – Extracting color values of semantic expressions," *IS&T CGIV*, Amsterdam, 355 – 361, 2012.
- [4] A. Lindner, B. Li, N. Bonnier and S. Süsstrunk, "A Large-Scale Multi-Lingual Color Thesaurus," *IS&T CIC*, Los Angeles, 30 – 35, 2012.