

## UNMANNED IMAGE NARRATION

SANGAVI VP<sup>1</sup>, SALEENA B<sup>2</sup><sup>1</sup>Department of Software Engineering, School of Computing Science and Engineering, VIT University, Chennai, Tamil Nadu, India. <sup>2</sup>School of Computing Science and Engineering, VIT University, Chennai, Tamil Nadu, India. Email: saleena.b@vit.ac.in

Received: 19 January 2017, Revised and Accepted: 20 February 2017

## ABSTRACT

Image analysis and understanding are used in various fields such as satellite imaging, robotic technologies, medical imaging, and defense. Image analysis is the process which involves extraction of meaningful information from images. Image annotation is the process of adding metadata to the image. The meaningful information is extracted based on the metadata available with the image. This research aims at understanding the input image and describing it to the user. The input image is segmented. The segments are annotated based on its properties such as color and shapes in it. The segment annotations are collected and processed to get the overall meaning of the image. This research would improve image search, inform the user about the image even if the image is not displayed due to compatibility issues, and it is unmanned – no human assistance required to understand and display the image contents.

**Keywords:** Annotation, Color Code, Semantics, Reasoning, Scenic Images

© 2017 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>) DOI: <http://dx.doi.org/10.22159/ajpcr.2017.v10s1.19534>

## INTRODUCTION

Image analysis is a domain which requires high degree extraction of information from images and it is a subject where the semantic gap impacts the user. Image analysis varies from simple to complex tasks, ranging from barcode scanning to face recognition. Image analysis involves computers which aid in the analysis of huge datasets, to handle the complex computations. This results in the extraction of relevant information. The human visual cortex is an instance apparatus for image analysis [1]. It aids to extract high-level information [2]. Many applications such as medicine, remote sensing, and security human involvement cannot be substituted by the computers. Human beings are important to image analysis tools and serve as inevitable resources for edge detectors and neural networks. These are motivated by human visual perceptions [3]. Image analysis contains the sections of machine vision, medical imaging, in which pattern recognition, signal processing, and digital geometry are used. The semantic gap speaks the difference between two descriptions of an object by linguistic representations. A semantic network is formed when several attributes related with one another, leads to build a domain. This is a bottom-up approach [4]. It evolves a computational representation.

A textual annotation is a tag that is added to the image or any other kind of content. It always depends on the knowledge, language of the annotator, and capability of expression. Hence, it could vary based on the attributes. To capture the displayed sceneries from the color of an image, algorithms are used for selection and manipulation of pixels. These algorithms are coupled and parameterized and mapped with the description. All the results on combining would display the contents of the image.

## RELATED WORK

In this paper “semantic image analysis for intelligent image retrieval” by Khodaskar and Ladhake [5], have used the knowledge-driven approach in performing semantic analysis. They started with image content analysis with respect to semantic concepts. Their work contained the database with images and required knowledge. The research was based on semantic content and image reference database for the knowledge delivery. A paper titled “a contribution to image semantic analysis,”

authored by Stašák [6] deal with the semantic analysis of an image. In his work, he divides the image into areas which he termed as segments, later each segment was linked to one or more semantic networks. These semantic networks were used to provide an image description. This was also used to complete images based on segments. The verbal facts helped to complete an image through segments. This specific paper helped to link segments with verbal tags leading to the annotation of image segments.

Later a paper by Aghbari and Makinouchi [7] titled “semantic approach to image database classification and retrieval,” presented an approach to hierarchically classify images based on the semantics. The semantics was constructed based on the perception of the image. This paper proposed to use support vector machines (SVMs) tool to classify images. Image segmentation into regions was performed through hill climbing method. This was more helpful for the work as it helped me in the semantic relation building process. The SVMs learned the semantics of specified classes; through a test data set of segmented image regions. It led to the implementation of intuitive query interface. A similar approach was found in “image semantic classification using SVM in image retrieval” by Yu and Liu [8]. They proposed a system which led to more flexible and intuitive query interface through human perception to learn the semantics of an image. The use of the binary SVM classifiers that classify image regions using different features at different levels of the hierarchy. This method helped greater precision of classification. Yanxi Liu *et al.* [9] in this paper “Classification Driven Semantic Based Medical Image Indexing and Retrieval” have coined a semantic-based image retrieval system that can discriminate between images through domain-specific cues. A similarity metric was computed based on image semantics. This similarity metric was used to image retrieval. The work also included analyzing colors and the color codes.

## PROPOSED SYSTEM

The proposed system aims at understanding the input image and describing it to the user. The input image is segmented. The segments are annotated based on its properties such as color and shapes in it. The segment annotations are collected and processed to get the overall meaning/contents of the image. Fig. 1 depicts the detailed flow of algorithm of the proposed system, and the further sections describe

in detail the phases in unmanned image narration such as image segmentation, image narration, and deriving semantics.

### Image segmentation

The input image is loaded from the specified path. The loaded image is then broken down into nine segments. The length and breadth of each segment is determined by estimating the length and breadth of the image and dividing it by the number of image segments to be made. The image segment is created and all the segments are stored in the form of the image array.

### Image annotation

The image segments are taken one by one, and the R, G, B color code of each pixel in the segment is taken to compute the average of overall R, G, B code of the segment. The difference between the average red, green, and blue is calculated, as this would bring down the overlap of colors. This process provides the major color of each segment by finding the greatest of all the average values. Based on the major color of the segment key is assigned to the segment. The key number maps the segment with the possible annotations. The possible annotations were derived based on the color and color code analysis.

### Deriving semantics

All the possible annotations of the specific segment combined with the location of the segment in the image will describe the segments. The same process is iterated across all the segments. For example, an annotation as "trees/grass" for the image segment in the location at the top (segment 0, 1, 2) of the image which leads to a single conclusion of the image segment to be annotated to "trees." All the descriptions together form the semantics of the whole image and hence displayed as the output.

### RESULTS AND DISCUSSION

The code was tested with random images to get the results based on its color properties and the location of image segments in the image. Annotation and reasoning were carried out during this process for finding the image contents. The input to the program (Fig. 2) is provided through a path by the user. The input image acquired is then broken down into nine segments (Fig. 3).

The color codes of the segments are determined and based on the major color of the segment, a key is assigned, and this key maps the possible annotations for the major color of the segment (Fig. 4).

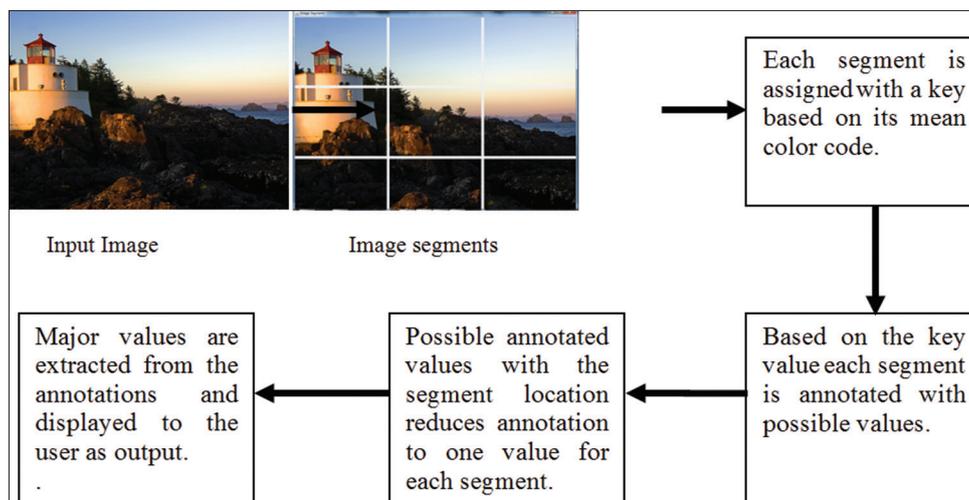


Fig. 1: Flow of algorithm



Fig. 2: Input image



Fig. 3: Image segments

```

<terminated> reasoning [Java Application] C:\Program Files\
Image:0 Color(key):5
Image:1 Color(key):5
Image:2 Color(key):5
Image:3 Color(key):1
Image:4 Color(key):5
Image:5 Color(key):5
Image:6 Color(key):0
Image:7 Color(key):0
Image:8 Color(key):0

POSSIBLE SEMANTICS OF THE IMAGE SEGMENT:
Image:0 Sky/Water
Image:1 Sky/Water
Image:2 Sky/Water
Image:3 Sunrise/Sunset
Image:4 Sky/Water
Image:5 Sky/Water
Image:6 Rocks/Hill

```

Fig. 4: Image segment keys

```

<terminated> reasoning [Java Application] C:\Program Files\
Image:3 Color(key):1
Image:4 Color(key):5
Image:5 Color(key):5
Image:6 Color(key):0
Image:7 Color(key):0
Image:8 Color(key):0

POSSIBLE SEMANTICS OF THE IMAGE SEGMENT:
Image:0 Sky/Water
Image:1 Sky/Water
Image:2 Sky/Water
Image:3 Sunrise/Sunset
Image:4 Sky/Water
Image:5 Sky/Water
Image:6 Rocks/Hill
Image:7 Rocks/Hill
Image:8 Rocks/Hill

CONCLUSION: The image contains
Sky
Sunrise Sunset
Rocks

```

Fig. 5: Annotations based on key values and conclusion

The last phase of the processing (Fig. 5) shows the output of the program which displays the description to the user. In the future, audio

function can be added to this project and this could be used to enable blind people to enjoy the experience of the image, in web applications

if the image is not displayed on browser the user can get the image contents automatically, the developer of the web page need not define the contents for every image.

### CONCLUSION

The contents of the image are identified based on the average color in the image segments and the location of the image segment in the image. Therefore, the image contents are derived based on the color scheme and the location of the image segment. This project has greater potential to reduce the gap between human perception and description of the image.

### ACKNOWLEDGMENT

I would like to thank the management of VIT University for giving us an opportunity to work on our idea and to develop it. I am thankful to Dr. B. Saleena for giving me an opportunity to work with her and do a publication with her help. I would like to express my gratitude to her for the invaluable support, help, and guidance.

### REFERENCES

1. Lee CH, Chiang KC. Latent semantic analysis for classifying scene images. Proceedings of International MultiConference of Engineers and Computer Scientists 2010. Vol. II. Hong Kong: March 17-19; 2010.
2. Cheng YC, Chen SY. Image classification using colour, texture, and regions. *Image Vis Comput* 2003;21:759-76.
3. Swain MJ, Ballard DH. Color indexing. *Int J Comput Vis* 1991;7(1):11-32.
4. Town CP, Sinclair D. Content Based Image Retrieval Using Semantic Visual Categories. Technical Report TR2000-14. Cambridge: AT&T Laboratories; 2000.
5. Khodaskar A, Ladhake S. Semantic image analysis for intelligent image retrieval. International Conference on Intelligent Computing, Communication and Convergence (ICCC-2014), Conference Organized by Interscience Institute of Management and Technology, Bhubaneswar, Odisha, India. *Proc Comput Sci* 2015;48:192-7.
6. Stašák J. A contribution to image semantic analysis. *INFORUM 2004: 10<sup>th</sup> Conference on Professional Information Resources*. Prague: May 25-27; 2004.
7. Aghbari Z, Makinouchi A. Semantic approach to image database classification and retrieval. *NII J* 2003;7:1-8.
8. Yu X, Liu H. Image semantic classification using SVM in image retrieval. (CD-ROM) Proceedings of the Second Symposium International Computer Science and Computational Technology (ISCST 09), Huangshan, P.R. China: 26-28, December; 2009. p. 458-461.
9. Liu Y, Dellaert F, Rothfus WE. Classification Driven Semantic Based Medical Image Indexing and Retrieval. *CMU-RI-TR-98-25*. Pittsburgh, PA: The Robotics Institute Carnegie Mellon University.