International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163
Volume 1 Issue 11 (November 2014) www.ijirae.com

# A Survey on Image Context Extraction Method

Pranay Tiwari
*Department of Computer Engineering, Sinhgad College of Engineering, Savitribai Phule Pune University, Pune*

***Abstract— In the recent years, Web has emerged as one of the most important source of information in the world. Images on the web consist of valuable contextual information that can be used for image annotation. This contextual information is nothing but the information related to or in context of image, which can be used for indexing the images. Therefore, to get qualitative descriptions of information, methods that are able to extract the information from image context become essential. This paper, discuss the sources in web documents where such useful information is located and different methods or techniques that can be used for extraction of contextual information.***

***Keywords— Web, Contextual Information, Image annotation, Web documents, Indexing***

## I. INTRODUCTION

Now-a-days, web has become an integral part of our lives comprising billions of web documents with information related to each and every conceivable topic. The information can be presented by different types of media, such as text, image, audio and video.

After text, image is the most well-known media type on the Web. However, as compared to text, the organization of images by means of the semantics they depict is much more complicated. While humans can easily recognize objects depicted in images, but in computer vision the automatic understanding of images is still one of the most challenging tasks. The main difficulty that arises in automatic understanding of images by computer is known as Semantic Gap [1]. Semantic gap is the lack of correlation between the semantic categories that a user requires and a low level feature that system offers. The semantic gap between the low-level visual features (color, shape, texture, etc.) and semantic concepts identified by the user remains a major problem in image retrieval systems. Therefore, bridging the gap between the low-level visual features (color, shape, texture, etc.) and semantic concepts identified by the user remains a major problem in image retrieval system.

Images on the web often come with the valuable textual information on hosting web pages that share similar semantics with the images. Such information is called as *contextual information.* Contextual information is the text that related to or in context of web image. This information is extracted from various sources in the web page that are related to the image. A major drawback in using such information is that it is subjective related to the authors point of knowledge, view and experience. But, irrespective to this flaw contextual information is rich with high –level semantic concepts and contains both direct and indirect information about the image. Therefore such information can be used to annotate the image and help in bridging the semantic gap between the image retrieval and the image semantics understandable by humans [2].

Even though the contextual information is very useful to us, but the main challenge lays in implementation of an extractor that carters to the assorted webpages for the definition of the surrounding text. This paper contains several sections that describe the locations in webpage from where we can get the desired information related to image and various methods for image context extraction. Section 2 describes the image context sources in web documents. Section 3 describes the challenges that web image context extractor faces and applications where this information can be used. Section 4 describes the methods for extraction of contextual information. Finally, Section 5 summarizes the paper.

## II. IMAGE CONTEXT SOURCES IN WEB DOCUMENTS

In earlier days, complete web page was used to extract the contextual information. This technique was inefficient as it used contain large amount of irrelevant text (i.e. information). There is no standard that can be used to relate text to an image in any existing image retrieval system. Only small part of the web page contains information that is relevant to image and not the complete web page. Therefore, it is important to know the sources in web page from where we can extract the information. In the following, we want to introduce and discuss the benefits and the shortcomings of possible image context sources.

A. ***Image URL***: It contains file name of the image and its path on the web server. If the file name is properly chosen it can include the keywords that describe the image content. As, some web designers organize web images in hierarchical folders e.g., “http://www.example.com/people/german/poet/Heinrich Heine.jpg” which allows to build taxonomy-based annotations derived from image URLs [3]. There are very rare URLs on the web that are well organized and there are hardly any web images whose filenames are manually assigned. Now-a-days, images possess generated filenames which are of no use.

***B. Page Title***: It is a short description of the web page content and therefore can prove to be an important hint to the image semantics. But, many web pages do not comprise of a single topic but diverse topics for example, start pages of most news domain. In such cases, the page title is hardly related to the images of that page, because it is either too general or contains information to the global web site, which in both cases would generate a misleading image description.

***C. Alternative Text (ALT)***: The value of ALT-attribute corresponding to the image tag was originally intended as a short image description to be displayed in environments where the image cannot be rendered. If properly used it can be the most useful source for image context, as it is directly related to the image and can be easily extracted. [4] Observed that only 21% images that they analyzed in their studies possess an ALT-attribute. Hence the benefit of ALT as image context provider is limited.

***D. Full Text***: As full text, complete displayable text of wen document is considered. Image and text content both can describe the similar information in different ways and therefore images and articles on web page must be semantically cohesive. Web documents are filled with contents having multiple topics and different images on the web page can belong to different particular content, but in this approach each image is associated with all the contents. The result of this approach consists of high recall, but precision is very low and thus misleading. Web pages contain several noisy elements such as navigational elements and layout, advertisement which do not contain any useful information and thus increase precision.

***E. Associated Text***: In this approach a small text portion of full text which is semantically related to the image content is selected. For example, some web pages have image caption with a small description of image context, which is similar to alternative text and hence may own the same quality properties. Even if the image caption is not explicitly present, the text surrounding the image is considered to share high degree of semantics with the image and thus can prove to be beneficial. [4], [5] and [6] showed that the associated text have ability to describe web images with more quality and reliability than other context sources. However, the associated text is not explicitly given therefore proper extraction methods are to be found.

## III. CHALLENGES AND APPLICATIONS

The extracted information should be such that it should benefit the application in it is which used. This section describes the challenges that are faced during the image context extraction and applications which are benefited by this information.

***A. Challenges:*** The detection of image and associated text oh hosting pages involves several challenges that are mainly set by the web:

- *Cluttered Web Pages.* Web documents are filled with multi-topic contents and the identification of a particular content block is unavoidable in order to guarantee precise image description. Moreover, with main content web page also consist of noisy contents such as advertisement, which makes extraction process more complicated.
- *Dynamic Contents.* Information and representation structure of the web page changes very quickly. It was observed that 40% of the data on the web changes every month. As layout structure of web page is also volatile, the extraction methods cannot rely on strict extraction rules since they have to be adapted whenever the layout changes.
- *Large Volume.* Large amount of data is available on the Web. One can find information about each and every topic on the web. But the rapid growth of the web poses scaling issues, which are not easy to cope with. Therefore the extraction method must be simple and fast.
- *Heterogeneous Data.* Information available on the web is heterogeneous. Contents are structured using different patterns among different web pages. Even, for a single webpage, the context can vary from a small caption to long article. The extraction method should be such that they can handle such heterogeneous structure.
- *Data Quality.* Everyone can publish their information on the Web while there is no such editorial institution that proves the correctness of the shared document. So, there is no guarantee on the quality of the shared contents. For example, we could find a web image with a dog and the surrounding text tells us about cat. This problem cannot be solved by extraction method and thus shows us the limitation of image context extraction approach.

Therefore, it is important to keep these problems in mind and find new solutions to the web image context extraction, as well as to realize the quality limits of image context in applications that could get benefited from this context.

***B. Applications:*** The context extracted from web page can be beneficial for many applications. Some application areas found in literature are;

- *Web Image Retrieval.* It is the most obvious application that uses image context. In this application the textual description associated with an image is used to generate an index for the image. Indexing can be done by using standard methods from text retrieval. User quires are matched against the indexed terms and the corresponding image is retrieved.
- *Web Image Annotation.* The aim of this application is to find the words or terms that describe the depicted objects best. It can be used as input to many applications. For example, it can be used in image retrieval system as described above; it can be used also in image organizing and image browsing applications. The learning of visual topics and concepts can also be done with the help of image annotation.
- *Learning of Visual Concept.* It describes the process of finding correlations between visual description of images and annotation. The image depicts the object described by keyword, if correlation occurs more often. In this way translation of concepts (i.e. visual to textual) are learned and can be stored in visual library. This library can be used for annotation of unlabelled image.
- *Web Image Redesign.* It is a process of restructuring a web page so that it can be seen on small embedded device such as mobile phones and PDAs as they have smaller displays. A web page can be displayed as a photo album with all content images, and then the user can chose the image from the list for the corresponding content.

The extracted contextual information can also be used to improve the quality of accessibility applications. When images do not have alternative text, then in such condition contextual captions can be of great benefit

## IV. EXTRACTION METHODS

Web pages are semi structured documents and therefore the context of images embedded in web pages is given only implicitly. There are three main approaches for image context extraction: *window-based approach, page segmentation* and *structure based wrappers*. The window based approach is a heuristic method that extracts text surrounding the image. Different variants of this approach have been made that are usually built on different assumptions in respect to length and surrounding text. In page segmentation approach web page is divided into segments of common topics and then each image is associated with the textual contents of the segment which it belongs to. The structure based wrappers are more adaptive heuristics rule that incorporate structural information form HTML codes of the web page in order to estimate borders of image context.

***A. Window Based Context Extraction:*** It is fast and most commonly used method for extraction of contextual information. In this subsection, N-terms Window algorithm [6, 7] is described along with its variant that uses shifted and dynamic window [4].

- *N-Terms Window.*

The idea behind this algorithm is to extract the text surrounding the image in HTML source code as contextual information. The algorithm of this approach is quite simple and easy, as shown:

***Algorithm: N-Terms Window***

**Data**: Web Document d, Image I, Window size N
**Result**: Set T of terms surrounding the image

```
1.  S ⟵ get Images and terms (d);
2.  i ⟵ index of(I,S);
3.  for k ⟵ [i- (N/2)] to [i+ (N/2)]
4.  |   if S[k]instances of term then
5.  |   |       T.add(S[k])
6.  |   end
7.  end
```

As input, we provide the web document, an image of this document and a window size *N*. In first step document is transferred into sequence of terms and images *S*. The terms are nothing but the particular words of the textual content of the given webpage and the images corresponds to the images in the document embedded by **<img>** tags. The elements in sequence *S* are ordered according to their position in original HTML code of the web document. Further, *i*-th element of *S* is denoted as *S[i]*.

In the next line (line2), the position of image present in sequence *S* is determined and stored as index *i*. This step is required to know the where to position the window frame in the next step.

The main part of the algorithm is between line3-line7 which is a loop that iterates over *S* from *S [i- (N/2)]* to *[i+ (N/2)]* and collects the terms that are visited. Provided that element *S[i]* is an image, the web image context is estimated as given in Fig 3.1.

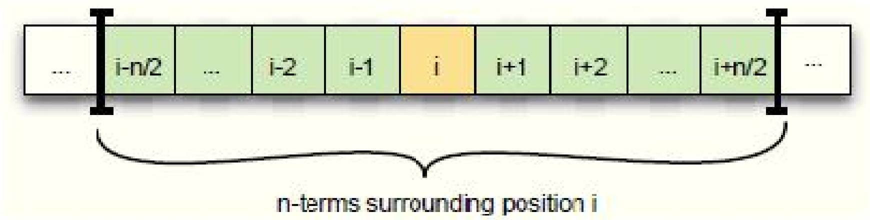



Fig 3.1: Frame of *n* terms surrounding an image in a list of terms and images.

Let the parameter *n* determines the size of the frame for term surrounding the images and has to be estimated. [6] have used different frame size in their evaluation and found that the value of $n = 20$ performed best. [7] have applied a frame for 30 terms, 10 terms before the image and 20 terms after the image were taken. [4] states that for the examined images 27% of the context appears before the image and 73% after the image.

The time complexity of the described algorithm is linear, since the image index estimation, the web page transformation and the window of terms computation are executed sequentially and each of them are linear in time as it depends on length of the document.

This approach also extracts terms that are irrelevant to image along with relevant image. Therefore, to better adapt the surrounding frame size, [8] has extended the basic method by leveraging HTML tag hints.

- *Adaptive Window.*

  The window based extraction technique mentioned above suffers from the fixed-length problem of the window and thus cannot deal with varying context sizes among webpages. To overcome this problem, [8] introduced a more dynamic method, which involves structural information contained in HTML. It was recognized that there are certain structural tags that are used by web designers in order to separate the contents. Some of the separator tags that **{<hr>, <br>, <p>, <table>, <tbody>, <tr> and <td>}**. These tags can be used as clues for context borders and better precision for the context can be estimated.

**Algorithm 2:** Adaptive N-Terms Window

```
   Data: Web document d, image I,
   Result: Set T of terms surrounding the image
1  S ← getTagsAndTerms(d);
2  i ← indexOf(I, S);
3  t = 0                          /* terms counter */;
4  k = i + 1              /* loop index for sequence S */;
5  while t < 32 do
6  |   if S[k] instanceOf TERM then
7  |   |   T.add(S[k]);
8  |   |   t = t + 1;
9  |   else
10 |   |   if S[k] is SEPARATOR then
11 |   |   |   terminate loop;
12 |   |   end
13 |   end
14 |   k = k + 1;
15 end
16 t = 0;
17 k = i − 1;
```

```
18 while t < 32 do
19 |   if S[k] instanceOf TERM then
20 |   |   T.add(S[k]);
21 |   |   t = t + 1;
22 |   else
23 |   |   if S[k] is SEPARATOR then
24 |   |   |   terminate loop;
25 |   |   end
26 |   end
27 |   k = k − 1;
28 end
```

The extraction algorithm needs an image and web document for input as the size of window is 64 at maximum. The first difference as compared to *n-term window* extractor is that all tags are additionally defined while initialization of the sequence *S*, tags are ordered in *S* by their position in HTML codes. In second line, index *i* of image *I* in Sequence *S* is estimated. Line 3-4 shows the initialization of the index and the terms counter for traversing *S*. The first while loop from line5-14 iterates over the element of sequence *S* that is after the image and checks whether they are terms or separator tags. If the current element is the term, the element is added to the context set *T*. The loop comes to an end, when the 32 terms are collected or either the element is a separator tag. The loop from line 17-28 is repeated in same manner for elements that are before the image in sequence *S*.

This method has been successfully applied by [4] and also [9] has applied a variant for this method in image annotation application.

Time complexity for the all the variants of this method remains linear, as the tags are included in sequence *S*, it depends on both number of tags and text elements.

**B.** ***Structure-based Wrappers:*** Images on the web are often used in news article, tagged albums, product catalog and similar type of environment. The similarity between all this applications is that the particular image is used as the part of the data record, which is usually stored in some database and displayed following some fixed template. Programs that are built to extract such structured data records from web pages are called ‘Wrappers’ [10].

Two wrapper-based approaches are mentioned below, which follow some manually created extraction rule that works for the most web images.

- *Paragraph Extractor.*

As the name implies, this wrappers aims at finding the nearest paragraph from an image and consider this paragraph as image caption. This is a *DOM* Based approach and to determine the context paragraph it uses parent child relation between *DOM* elements by estimating the parent tag element of the given image element, which include text elements in its subtree. All text elements under the estimated parent tag are considered as parts of the image context.

***Algorithm 3: N-Terms Window***

**Data**: Web Document d, Image I, Window size N
**Result**: Set T of text node representing the image caption.
***1.*** *D⟵ Create DOM(d);*
***2.*** *i ⟵ Find (I,D);*
***3.*** **while** *contains Texnode(i)* **do**
***4.*** | *i ⟵i.getParent;*
***5.*** **end**
***6.*** *T⟵ getTexnodes(i);*

In the initialization step line1-2, the image element *I* corresponding to the input image and the DOM tree *D* are computed. The while loop (lines3-5), which starting at the image element, walks the tree upwards, until a parent element is reached that includes text elements in its subtree. Line 6 is the final step of the algorithm where the text elements under the determined parents are collected as image context.

Time complexity of the algorithm is directly linear to the length of the document in regards to building the DOM tree and further the while-loop is maximum at *O(d.log(t))*, where *d* represents the maximum degree of the DOM tree and *t* as its depth.

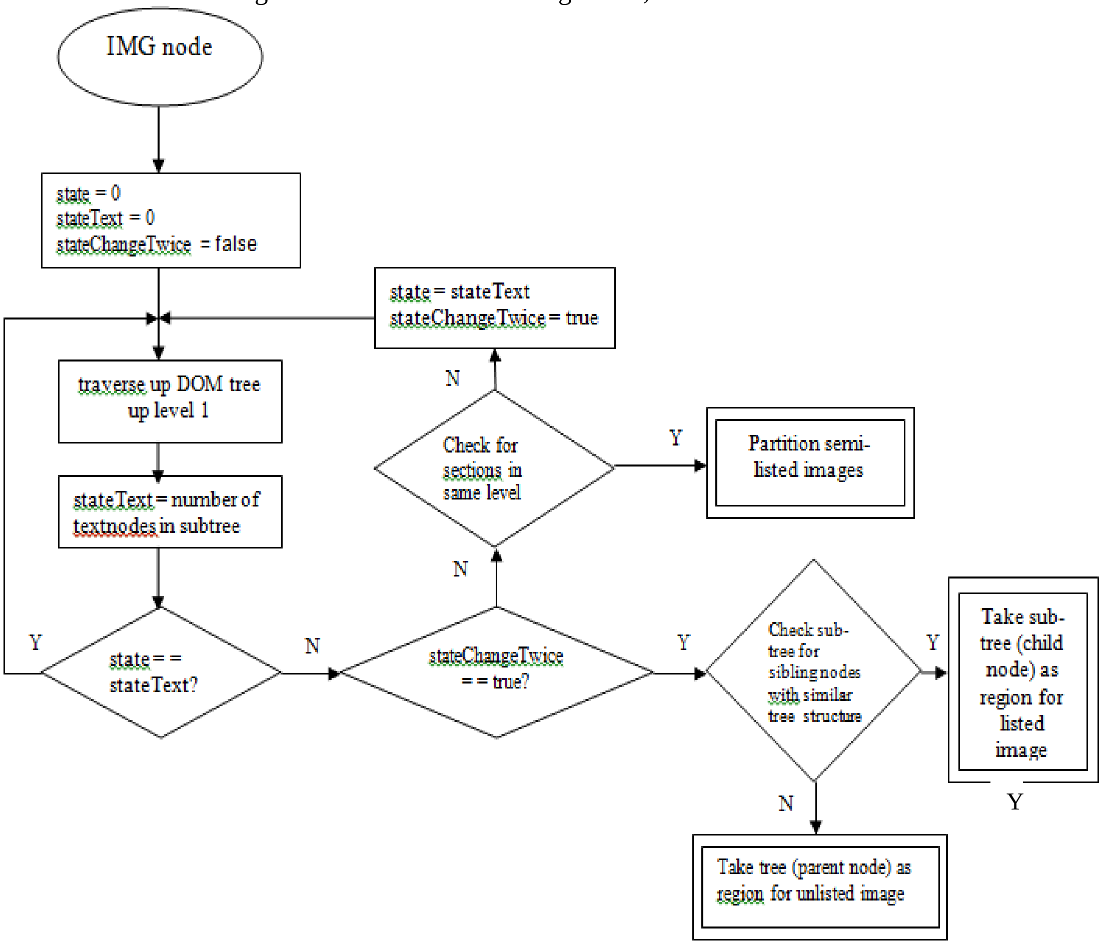
- *Monash Extractor.*

The Monash Extractor is an extension to the Paragraph Extractor and has been introduced by [11] to handle the different template types in which an image can be embedded in. The basic idea on which the method relies is the concept of detail pages and list pages [10] in which generally extractable data records can be placed in. Detail page contain detailed information of one particular record, while list pages contain list of many records with similar structure.

The algorithm proceeds as, DOM tree and image node whose context is to be identified are taken as input. Three state variables are used that maintains the current state of the algorithm: **state** and **stateText** are the two variables that keep the current number of text nodes under the actual nodes and its previous value. Both this variables are set to 0 at beginning. The third variable **stateChangeTwice** is true if the state variable has changed two times during the actual run.

Starting from the image node, the algorithm walks upward the DOM tree, until the number of nodes under the tree has changed. The algorithm now checks whether the number of text contents has changed twice. In the first run, the current node is checked for typical semi-listed structure (in semi-listed structure DOM tree representation, the segment of particular image are not placed under one root node, but they are all together under a root node while the visual separation are made by HTML tags) with repeating patterns of HTML tags. If such pattern is found while traversing then it is considered as semi-listed image and the region of the image is extracted and returned as image context. If no semi-listed structure is found, the algorithm continues to traverse the tree upwards until a parent node is found at which the number of nodes in the sub-tree changes again. In such a case, the number of nodes changes twice and the subtree is checked for siblings with similar tree structure. If such sibling nodes are found, the current image is and the subtree which belongs to the image is returned as image region. Otherwise, the image is considered as unlisted image, and the complete tree under the actual node are returned as image region.

The following flowcharts describes the algorithm;



The time complexity of this method is same as Pragraph Extractor, since the text node and DOM traversing are also both conatined in Monash Extractor.

C. ***Context Extraction by Page Segmentation:*** As stated above, the main content of the web page consist of multiple topics. Beside, main content it also consist of additional elements such as navigation and advertisement bar. Web page segmentation, describes a task of partitioning a web page into disjoint blocks of coherent content.

Web page segmentation is used as preprocessing in many application area such as, content de-duplication [12,13], keyword-based search [14], web page reformatting for displaying on small scale device [15] and web image annotation [16,17]. Two general approaches of web page segmentation are described below, which are characterized by their partition method: *bottom-up* and *top-down* approach.

- *Top-Down Page Segmentation.*

This approach starts by considering complete webpage as a block and iteratively dividing this block into smaller blocks using different features obtained from the content of the webpage.

The *Vision based page segmentation (VIPS)* algorithm [18] is used to extract the web image context. *Degree-of-Coherence (DoC)* is computed for each block by VIPS utilizing heuristic rule based on DOM as well as visual features. The range of DoC ranges from 1-10, where 10 denote maximum value of coherence. A *Permitted Degree of Coherence (PDoC)* has to be defined by user that stands for the maximum DoC at which portioning is performed. In order to control the granularity of partitioning different values of PDoC are used. VIPS compute *Visual Content Structure Tree* as a by-product that is assumed to possess a more semantic separation of content as compared to the traditional DOM tree.

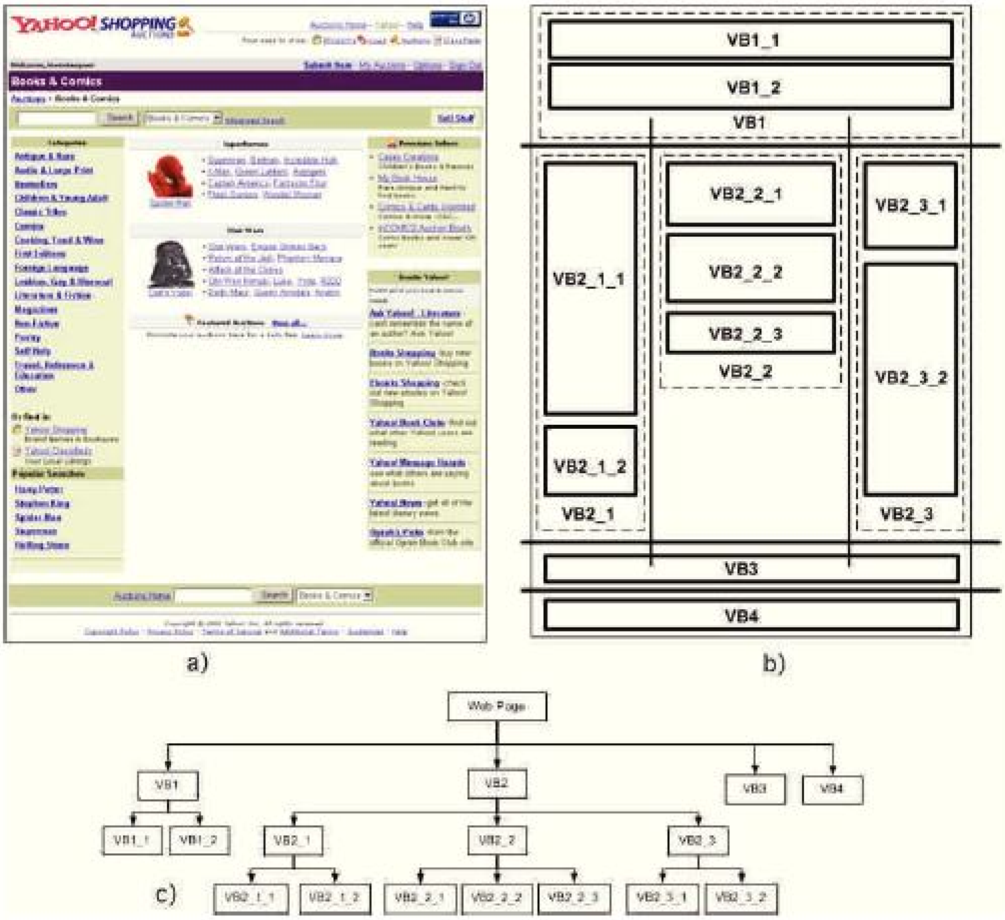



Fig 3.3: Yahoo! Auction web page (a) in browser view, (b) as block representation, (c) as tree representation.

[19] combine two different methods for web page segmentation. The first method applies heuristic rule that is based on HTML tags. In the second method, based on the order of HTML tags a content-distance is defined and where the content-distance exceeds a dynamically estimated threshold at such a position the initial block is iteratively separated.

[20] defines a content size and an entropy value that measures the strength of local

patterns within the sub tree of a node. To perform page segmentation threshold values are to be defined for both measures.

[15] defines a segmentation system, which divides a webpage into nine segments using a decision tree which employs an information gain measure and geometric features.

- *Bottom-Up Page Segmentation.*

In the bottom-up approach of web page segmentation, the leaf nodes of the DOM representation are taken as atomic content units mostly. These units are grouped into segments in the next step.

[12] approached a problem that arises during page segmentation from a graph-theoretic point of view. The edge weights estimate the costs needed to put the connected nodes in one block and DOM nodes are supposed to be nodes in complete weighted graph. This technique is used to reduce the partition task to well-known optimization problem. A solution to the problem is given by the Energy Minimizing Graph Cuts algorithm. As features, visual as well as DOM based features are applied.

[13] represent atomic content unit of a webpage by a quantitative linguistic measure of text density and reduce the segmentation problem to solving 1D-partition task. An iterative block fusion block applying methods adapted from computer vision is presented as a solution.

## V. CONCLUSIONS

In this paper, detailed literature about various web image context extraction techniques is mentioned. The sources on web page from where we can extract the contextual information are also mentioned along with the challenges that occur while extracting information and the applications where the extracted information can be used. This various extraction techniques have their own advantages and disadvantages. However, there are different documents, like PDF or open office files that similarly include images with corresponding context and it would be interesting to see that proposed algorithms could also be applied on such documents. Another direction of future work are the applications that make use of the web image context.

## ACKNOWLEDGMENT

Author likes to thanks ***Prof.P.M.Kamde*** for their valuable guidance in preparing this paper. Author also likes to thank ***Prof. C.A.Laulkar*** and ***Prof. P.R.Futane*** for their support in this work and also to the unseen authors of various articles on the internet helping author to become aware of the research currently ongoing in this field.

## REFERENCES

[1] Hui Hui Wang, Dzulkifli Mohamad, N.A. Ismail, “Approaches, Challenges and Future Direction of Image Retrieval System,” *JOURNAL OF COMPUTING*, volume 2, issue 6, JUNE 2010, ISSN 2151-9617.
[2] Chee Sheen Chan, Adel Johar, Jer Lang Hong, “Contextual Information for Image Retrieval Systems,” in 10th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2013.
[3] John R. Smith and Shih-Fu Chang. *Visually searching the web for content*, IEEE Multimedia, 4:12–20, July 1997.
[4] Huamin Feng, Rui Shi, and Tat-Seng Chua. “A bootstrapping framework for annotating and retrieving WWW images”. In Proceedings of the 12th annual ACM international conference on Multimedia, MULTIMEDIA’04, pages 960–967, New York, NY, USA, 2004. ACM.
[5] Sadet Alcic. “Improvement of Content-based Image Retrieval Quality using Automatic Annotation”, April 2007. Master Thesis.
[6] Tatiana Almeida Souza Coelho, P´avel Pereira Calado, Lamarque Vieira Souza, Berthier Ribeiro-Neto, and Richard Muntz. *Image Retrieval Using Multiple Evidence Ranking*. IEEE Transactions on Knowledge and Data Engineering, 16(4):408–417, 2004.
[7] Stan Sclaroff, Marco La Cascia, and Saratendu Sethi. “Unifying textual and visual cues for content-based image retrieval on the World Wide Web”. Computer Vision and Image Understanding, 75(1-2):86–98, 1999.
[8] Lexin Pan. *Image8: an image search engine for the internet*. Honours year project report, School of Computing, National University of Singapore, April. 2003.
[9] Hua Zhigang, Wang Xiang-Jun, Liu Qingshan, and Lu Hanqing. “Semantic knowledge extraction and annotation for web images”. In MULTIMEDIA’05: Proceedings of the 13th annual ACM international conference on Multimedia, pages 467–470, New York, NY, USA, 2005. ACM.
[10] Bing Liu. *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*. Data-Centric Systems and Applications. Springer, 2007.
[11] Fariza Fauzi, Jer-Lang Hong, and Mohammed Belkhatir. *Webpage segmentation for extracting images and their surrounding contextual information*. In ACM Multimedia, pages 649–652, 2009.

[12] Deepayan Chakrabarti, Ravi Kumar, and Kunal Punera. "A graphtheoretic approach to webpage segmentation". In Proceeding of the 17th international conference on World Wide Web, WWW '08, pages 377– 386, New York, USA, 2008.

[13] Christian Kohlsch¨utter and Wolfgang Nejdl. "A densitometry approach to web page segmentation". In Proceeding of the 17th ACM conference on Information and knowledge management, CIKM '08, pages 1173–1182, New York, USA,2008.ACM.

[14] David Fernandes, Edleno S. de Moura, Berthier Ribeiro-Neto, Altigran S. da Silva, and Marcos Andr´e Gon¸calves. *Computing block importance for searching on web sites*. In Proceedings of the sixteenth ACM conference on Conference on information and knowledge management, CIKM '07, pages 165–174, New York, NY, USA, 2007. ACM.

[15] Shumeet Baluja. "Browsing on small screens: recasting web-page segmentation into an efficient machine learning framework". In Proceedings of the 15th international conference on World Wide Web, WWW '06, pages 33–42, New York, NY, USA, 2006. ACM.

[16] Xiaofei He, Deng Cai, Ji-Rong Wen, Wei-Ying Ma, and Hong-Jiang Zhang. "Clustering and searching www images using link and page layout analysis". ACM Trans. Multimedia Comput. Commun. Appl., 3(2):10, 2007

[17] Xiaoguang Rui, Mingjing Li, Zhiwei Li, Wei-Ying Ma, and Nenghai Yu. "Bipartite graph reinforcement model for web image annotation". In *Proceedings of the 15th international conference on Multimedia*, MULTIMEDIA'07, pages 585–594, New York, NY, USA, 2007. ACM.

[18] Deng Cai, Shipeng Yu, Ji-Rong Wen, and Wei-Ying Ma. *VIPS: a Vision based Page Segmentation Algorithm*. Technical report, Microsoft Research (MSR-TR-2003-79), 2003.

[19] Gen Hattori, Keiichiro Hoashi, Kazunori Matsumoto, and Fumiaki Sugaya. "Robust web page segmentation for mobile terminal using contentdistances and page layout information". In Proceedings of the 16th international conference on World Wide Web, WWW '07, pages 361–370, New York, NY, USA, 2007. ACM.

[20] G. Vineel. Web page DOM node characterization and its application to page segmentation. In Internet Multimedia Services Architecture and Applications (IMSAA), 2009 IEEE International Conference on, pages 1 –6, 2009.

[21] Parag Mulendra Joshi, Sam Liu, "Web Document Text and Images Extraction using DOM Analysis and Natural Language Processing", HP Laboratories,HPL-2009-187.

[22] Himali Chaudhari, Prof D.D.Patil, "A Survey on Automatic Annotation and Annotation Based Image Retrieval," International Journal of Computer Science and Information Technologies, Vol. 5 (2), 2014, 1368-1371.

[23] Shalini K.Kharkate, Nitin J.Janwe, *Automatic Image Annotation: A Review*, The International Journal of Computer Science & Applications (TIJCSA), Volume 1, No. 12, February 2013 ISSN – 2278-1080.