# ADAPTING WEB IMAGES FOR BLIND PEOPLE

A. Fatzilah Misman
*Department of Information System, International Islamic University Malaysia, Kuala Lumpur, Malaysia*

Peter Blanchfield
*School of Computer Science, University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham, U.K.*

Keywords: Web accessibility, Image tagging, Describing web image, Semantic web.

Abstract: One way to remedy the gap that evidently exists between the image element on the web and the web user who is visually blind is by redefining connection between the image and the abundant element of the web itself i.e. text. Studies on the exploitation are done largely related to the fields like the HCI, the semantic web, the information retrieval or even a new hybrid approach. However, often many see the problem from the perspective of the third party. This position paper posits that the problem can also be seen from the fundamental reasons for an image being on a web page without neglecting the connection that develops from the web user's perspective. Effective and appropriate image tagging may consider this view.

## 1 INTRODUCTION

Adapting web images to make them useful to blind users of the Internet presents a major technological challenge. In practice web images are commonly ignored by blind users due to the considerable difficulty in obtaining meaningful output. Evidence for this has been found from a preliminary study with users of a local blind centre. This is also supported by evidences from related studies (Petrie, O'Neill & Colwell, 2002, Petrie & Kheir, 2007).

However, "Meaning can be as important as usability in the design of technology (Shinohara & Tenenberg, 2009)". Thus theoretically the 'meaning' of an image is bounded by values of information an image can offer. King, Evans, and Blenkhorn (2004) indicated that there are four common technological solutions for blind users to access the Internet including screen reader technology working alongside a human assistant to interpret images for them. When using a screen reader most users will have to listen to the whole of an image tag and try to interpret the meaning of the image from the content of the tag. It is seldom possible for them to gain the same level of meaning from an image by this means as would be possible for a sighted user looking at the image. If an "alt" tag has been used the system may choose to read this alone. However, the purpose of an image is often not well defined by the content of these "alt" tags (Bigham, Kaminsky, Ladner & Danielsson, 2006, Petrie, Harrison & Dev, 2005).

The ability to re-tag images with appropriate descriptions is the main goal of this paper. There have been various other attempts to retag images but this paper proposes a novel approach to this retagging which deliberately emphasises the actual use which blind users will make of the images rather than merely trying to retag on the basis of inferred knowledge from the tags. For example King et al. (2004) demonstrated this by developing a dedicated web browser for blind users. They suggested that images can be ignored entirely except when they contain valuable information in the "alt" tags or when containing a hypertext link which gives meaningful information with regard to a hypertext destination. However, in the current research an ethnographic study has found that blind users will want to retain images if they have direct relevance to the content. Thus, we believe, if the image tag is exploited effectively there are circumstances where they can represent better information when read using a common screen reader. Once the relevance of an image has been established the blind user will then often gain the help of a sighted user to explain the purpose of the image.


Fatzilah Misman A. and Blanchfield P..
ADAPTING WEB IMAGES FOR BLIND PEOPLE.
DOI: 10.5220/0003402004300437
In *Proceedings of the 7th International Conference on Web Information Systems and Technologies* (WEBIST-2011), pages 430-437
ISBN: 978-989-8425-51-5

There are a vast number of images on the World Wide Web today. However, the standard guidelines for providing appropriate descriptions for them are not enforceable (Caldwell, Cooper, Reid, & Vanderheiden, 2008). Despite numerous studies on visual processing of images to produce annotation or explanation, there is as yet no software that can determine image content in a widely useful way (Russell, Torralba, Murphy, & Freeman, 2005, Li & Wang, 2008). In addition to image analysis approaches a number of other studies have been undertaken that make use of data mining principles specifically to produce good image tagging for images (Bigham et al., 2006, Bigham, 2007) The position of the current study is to establish a definition of a "meaningful image" for blind people and thus enable valuable retagging of images to take place. It embodies expectations expressed by blind people and exploits the role an image plays on the page it is in.

Current techniques for generating image descriptions based on text content are insufficient in many ways. It is generally assumed that elements in a web content (particularly body text and images that appear within the same page or within the same virtual boundary such as a title or a head line) are associated to one another. This presumption has however always been challenged – see for example Carson and Ogle (1996). They asserted that even at that stage images might not be related to the content of a page. For example an advertisement image may have no direct relationship at all to the content of a page. This concern may be more significant for issues dealing with information retrieval but may not be as important when dealing with providing access to information on images on a page for blind users.

## 2 RELATED WORK

Previous studies on generating image labels or descriptions have made use of text descriptions via the html's "longdesc" or "alt" syntaxes for image descriptions. Despite the significant amount of work that has been done the only reliable method available for providing precise description to this day is manual labelling (Ahn & Dabbish, 2004). The reasons are that this method is independent, unrestricted, self manageable and original especially when it is performed during the development phase when the author initially is adding image elements to the web site. The problem however is that this approach may become tedious and thus extremely costly, particularly when the web page has already been published.

As an alternative solution, studies on image description have looked into automatic approaches to labelling. Ahn and Dabbish (2004) developed a word-image-matching online game that exploited players contributions to generate descriptions for web images. The game involves pairs of players who are randomly partnered to guess what each other have keyed in for the same image. Every time both players describe an image similarly they can then move on to the next image and the word or phrase used is chosen to describe the given image. This method adapted the concept implemented in the "Open Mind Initiative" project (Stork, 1999, Stork & Lam, 2000). The Open Mind Initiative applied the intelligent agent concept where the machine is trained about the attributes of various sets of image groups derived from a training database which was also contributed to by public users. Outsourcing data from open contribution conveniently creates sufficiently representative training databases (Datta, Li & Wang, 2005). However, use of outsourced labelling fails to address fundamental issues of web image dynamism. For example, images from a web source can be interpreted differently by different people. Also, for example with news web sites, images are added to sites frequently and in large numbers. In addition a huge number of new pages are added to the Web daily.

However, automatic techniques which have built-in mechanisms to generate image descriptions are also under development. For example the ALIPR (Automatic Linguistic Indexing of Pictures – Real Time) used a large database consisting of web pages as a training resource to develop learning algorithm of image signatures (Li & Wang, 2008). The learned patterns cluster images discretely and categorise from binary text to image pixels. The clustered patterns are then used to produce real time image descriptions. However, dependence on a large database system would introduce the issue of reliability of the approach when running on sizable data.

Bigham et al. (2006) use a web domain that runs a combination of three methods, enhanced web context labelling, Optical Character Recognition (OCR) and human labelling, to perform the labelling task. The use of three independent systems has the effect of increasing processing load. The web context labelling utilises summaries of the page, the title or header to imply the image content. This system also uses the content of linked pages for the same purpose.

The Latent Semantic Analysis (LSA) was originally used to reduce the semantic dimensionality of information retrieval problems such as the use of the synonym concept to elaborate an object query (Deewester et al., 1990). The theory of the LSA has been widely used and has evolved into varieties such as the Explicit Semantic Analysis (ESA) by Gabrilovich and Markovitch (2007). The ESA particularly computes semantic relatedness among words of human natural languages by comparing the corresponding vectors using conventional metrics (e.g., cosine). The ESA experiment on Wikipedia®-base knowledge data demonstrated substantial improvement based on the yield assessment. It also demonstrated that, limited words such as the information tagged to an image could yield more relevant words which can thus be used to describe the image. The problem however, is that this approach is not suitable when small databases have to be used and so new approaches must be found.



# 3 PROPOSED APPROACH

The proposed approach has been influenced by the results of an ethnographic study as mentioned in the introduction. This study looked at the use being made of images by blind people when using the web. The study took place in a local centre for the blind, which the users attend to gain help in accessing the web and its content. It was clear from this work that general use of the internet by the blind was for similar purposes to that of sighted people (Shinohara & Tenenberg, 2007). The centre provided screen readers and the addition of sighted assistants for helping the blind users. Observations were made through the use of general conversations and later from direct questioning about the value or otherwise of image content. By understanding the real scenario and the expectation of blind people from web images they come across, the question arose as to whether or not images could be adapted to become meaningful for their benefit using the existing independent screen reader software such as JAWS® and SuperNova®. The responses were positive provided web images had readable descriptions that at least met one of the following criteria:

- Conciseness – that is the tag is concise and gives an accurate reflection of its relationship to the page content.
- Appeal – the image tag invites the reader to continue to read the page content.
- Readability - any understandable description carries value.

Specifically these ideas were derived from the experiences in the ethnographic study. During this process two approaches were adopted. In the first the sighted helper became directly involved in the browsing process, responding to requests from the blind person while they listened to the screen reader. When images were viewed their relevance would be assessed by the helper. If the tagging was too long or inappropriate the helper would provide their own interpretation. In the second the browsing is directed by the blind user. The helper would correct them if they went wrong in the process. The blind user would then ask help when they noticed a word that seemed relevant in the screen reader output. They would then request assistance from the helper. Often the blind user would be misinterpreting keywords in an image tag. However, they would mostly skip most of the image elements. A similar observation with a single user was made by Shinohara and Tenenberg (2007).

This paper proposes an approach that combines i) Image purpose ranks developed from the work of Paek and Smith (1998) on image purposes ii) an interpretation engine which filters and replaces the HTML within the image tag with the relevant semantic direct interpretation and iii) a Proximity Weighting Model (PWM) that weighs replacement tags by measurement of relevant phrases or sentences within close proximity (that is, page proximity). All of these approaches have been used before but in this case the result which is delivered to the blind user is interpreted with respect to the analysis of purpose results derived from the work with blind users. So instead of replacing an image with text, for example, its value to a blind user is first assessed. If the image is content related (within the definitions given in Table 1 below) then it will be retained but retagged, clearly and concisely indicating it is content. If it is purely navigational the tag should be adapted to reflect its navigational purpose but the image information (contained for example in the '<*img*>' tag) might be removed. This paper proposes that text within a web page retains a certain level of association to other elements within the page in various ways. Thus, adapting images for blind user means to generate a description tag which fulfils at least one if not all the three criteria stated above.

## 3.1 Image Purpose Rank

According to (Sutton & Staw, 1995) the 'core'

question for an explanation is the question 'why'. It equips an occurrence with a sense of reason. Paek and Smith (1998) suggested a technique to improve the cataloguing and indexing of web images based upon image contents. They identified image purposes associated to image content. They suggested that every embedded web image could be categorised as content, navigation, decoration, logo, advertisement, information and content as described in Table 1.

Table 1: Adapted from Paek and Smith (1998) purposes of web image.

| Purpose | Definitions |
|---|---|
| *Advertisement* | Image that contributes to the act of informing, notifying or promoting what it represents which may or may not be related to the page's content. (e.g. corporate logo, individual brand on an anonymous website) |
| *Decoration* | Image that is meant for decorating the page. (e.g. buttons, balls, rules, masthead, background) |
| *Information* | Image that signifies the message which it represents. (e.g warning signs, under construction, what's new, graph figures) |
| *Logo* | Graphic representation or symbol of a company name, trademark, abbreviation, etc., often uniquely designed for ready recognition. (e.g. IBM, corporate logo) |
| *Navigation* | Image that represents the act or process of navigating within or without the page. (e.g. hyperlinked image, arrows, home, image map) |
| *Content* | Image that is associated with a body of text of a source page. (e.g, Honda advert on Honda homepage, image that is transcribed in the page content) |

In order to verify that the current authors' interpretation of this claim was correct an online survey with a group of expert users whose professional experiences are related to Information Technology was made. Fifteen respondents took part where each assessed a randomly selected and reconstructed web page from the original source. Six images from each page were indexed with the respective question number. Each question states seven options to range from the six image purposes and one 'uncertain' option. The results of this survey were convincing that this interpretation was correct and so could be used for the classification process. Sixty three percent of the responses exactly matched those of the authors. Of the other responses there was a range of disagreement between the correspondents but the majority were generally more in agreement with the authors.

Next, image purpose profiling was developed to define the image classification categories. The classifications were developed based on a study of patterns derived from a total of 1707 images from 175 web pages chosen using Yahoo!® API. The websites were from fairly mixed domains consisting of organization, government, commerce, social networking and educational websites. The classification defines regular patterns of occurrences within image tags identified by keywords, HTML syntax and frequency of occurrence of each keyword. The identified set of keyword patterns were then considered as a set of definition criteria representing the respective image's purpose. Every keyword is given a weight that indicates its contribution to an image's purpose's on a web site. For example, the keyword 'ads' has a mean score of 74% matching to the advertising category and the keyword 'icon' has a mean score of 78% indicating it as a decorative image.

Therefore, the purpose ranking role is to generate semantic expression which indicates ranked purposes definition based on the ranking perspectives algorithm. Figure 1 shows an example result of semantic expression of purpose using this ranking.

Figure 1: Example of semantically expressed purpose rank- "This image is *quite likely* navigational and a content's *less likely* an advertising or a decoration or *the least* a logo or an information".

The algorithm basically works by calculating the accumulated mean weight of any identified pattern found within the image tag and dividing it by the number of pattern occurrences (see Table 2).

## 3.2 Semantic Interpretation of an Image Tag

While HTML image taglines carried by image function tags '*<img/>*', '*<map>*' and '*<area/>*'

Table 2: Semantic rank on purpose point score.

| Semantic rank | *More likely* | *Quite likely* | *Less likely* | *The least* |
|---|---|---|---|---|
| Mean score rank | 75 – 99 (%) | 50 – 74 (%) | 25 – 49 (%) | 1 – 24 (%) |
| Mean scores of Figure 1. | | Navigat = 0.6 Content's = 0.56 | Advert = 0.4 Decor = 0.3 | Logo = 0.2 Info = 0.2 |

have the intended interpretation for web developer's purposes their interpretation by web users would be far less obvious. However, blind users using a screen reader will encounter the whole of the image tag in place of the image. The aim of the current work is to replace these tags where possible with tags that more concisely represent the purpose of the tag. Thus as said earlier if a tag is merely navigational it can be replaced by an appropriately tagged link. However if it is content related it must get a suitably concise, appealing and readable tag.

Table 3: Example of image tag attributes.

| Required Attributes | | |
|---|---|---|
| **Attribute** | **Value** | **Description** |
| *alt* | Text | Specifies an alternate text for an image |
| *src* | URL | Specifies the URL source of an image |
| Optional Attributes | | |
| *height* | Pixels% | Specifies the height of an image |
| *longdesc* | URL | Specifies the URL to a document that contains a long description of an image |
| *usemap* | #mapname | Specifies an image as a client-side image-map |
| *width* | Pixels % | Specifies the width of an image |

The original image tag attributes provide information from which the interpretation of the image's purposes can be inferred and their acceptable values can be determined. The image tag attributes are structured and designed for the developer's convenience. This feature is particularly interesting for the case of blind users whereby image tags can possibly be used for semantic interpretation answering three questions related to the image. They are;

- What does the image do?
- Where does it come?
- What is the image possibly about?

Possible viable answers are contained within '*href=*', '*src=*' and '*alt=*' as presented in Table 3.

Table 4: Examples of image tag interpreted attribute.

| Img tag 1 | <img src="penguin-pictures-2.jpg" alt="penguin pictures" width="775" height="271"> |
|---|---|
| *Interpret 1* | *image from penguin pictures jpg file .noted as penguin pictures* |
| Img tag 2 | <a href="http://www.penguin-pictures.net/adelie_penguin_pictures.html "> <img src="adeliepenguinpictures.gif" alt="adelie penguin pictures" width="165" height="30" target= "blank"></a> |
| *Interpret 2* | *navigates to web site penguin pictures network adelie web page .image from adeliepenguinpictures gif file .noted as adelie penguin pictures.on new window* |

## 3.3 Proximity Weighting Models (PWM)

PWM is often applied in the first stage of the Latent Semantic Analysis (LSA), as part of the local weight function (LWF) (Nakov, Popova & Mateev, 2001). It enhances document retrieval effectiveness by examining documents in terms of local level statistics (Macdonald & Ounis, 2010). Modern information retrieval systems generally take more than a single-term weighting model to rank documents. The PWM has contributed a significant factor in addition to LSA assessment for context relatedness within a document (Croft, Metzler & Strohman, 2009). It is based on the theory which states that proximity or distance between words retains certain weight of connection. For example, a repetition of the same noun though positioned in different lines or in different paragraphs of a document, suggests that the word is important to those lines or to the paragraphs.

The objective of PWM application in this case is to measure relatedness of page content text by measuring their weight distance from the image. The weighting scheme and computation is adopted from the model presented by (Cascia, Sethi & Sclaroff, 1998, Sclaroff, Cascia & Sethi, 1999). This was selected because of the implication of the likelihood of useful information occurrences that may be connected to an image tag. Besides, this method is sensitive to the distance of an individual pair depending on the way aggregation occurs locally (Tao & Zhai, 2007).

To obtain result quality, preprocessing of the page document is performed i.e. removing stop

words and stemming. Each image's associated HTML tag is parsed and a word frequency histogram is computed such as in (Table 5).

Table 5: Terms from image tag. Cell entries are the number of times that words (column) appeared in the content's HTML tags (row).

tag_no_3

| TagFeat\NavImg_kW | penguin | pictures | adelie | adeliepenguinpictures |
|---|---|---|---|---|
| cntntxt_freq | 24 | 21 | 1 | 0 |
| <_description> | 2 | 1 | 0 | 0 |
| <_keywords> | 2 | 1 | 1 | 0 |
| <_title> | 2 | 1 | 0 | 0 |
| <_small> | 2 | 1 | 0 | 0 |
| <_strong> | 3 | 3 | 0 | 0 |
| <_em> | 2 | 2 | 0 | 0 |

Phrases containing associated words are not always similar in length and structure. Words appearing with specific HTML tags are considered special thus being assigned with higher weight than all other words - see Table 6.

Table 6: Cascia et al.'s (1998) word weight scheme on HTML tags.

| HTML Tags | Weights |
|---|---|
| ALT field of IMG | 6.00 |
| TITLE | 5.00 |
| H1 | 4.00 |
| H2 | 3.60 |
| H3 | 3.35 |
| H4 | 2.40 |
| H5 | 2.30 |
| H6 | 2.20 |
| B | 3.00 |
| EM | 2.70 |
| I | 2.70 |
| STRONG | 2.50 |
| < No Tag > | 1.00 |

These weight values are heuristically derived based on selective weighting of words appearing between various HTML tags that measure their estimated useful information implied to the text. The importance given to a word is computed from its frequency, location relative to tags as shown in table 6 and relative to its proximity to the image (as described below). For example, if the word 'adelie' appears within '<*bold*>' tags in the HTML once, then it is initially weighted at 3.0. If instead it appears in text without HTML tags it is initially weighted as 1.0. Multiple occurrences also carry multiplier to the single value.

For words appearing before and after a particular image tag, the proximity weighting value is computed from equation (1) where *pos* is the position of the word with respect to the image tag and *dist* is the maximum number of words considered in applying such weighting. In this implementation, dist is 10 for words appearing before the image and 20 for words appearing after the image. The constant $\rho = 5.0$ is assigned so that the nearest text to the image tag is slightly less than and equal to the words appearing in the alt field of that image and the 'title' of the page respectively.

$$\rho \cdot e^{-2.0 \bullet pos/dist} \quad (1)$$

This weight distance proximity function is based on the assumption that words close to an image have connection to the image. It is also assumed that the degree of the connection decreases exponentially with the distance of the recurring word from the original word in the image although this may not always be the case. This assumption is founded in reference to the analytical survey results made by Cascia et al. (1998) themselves and by Sclaroff et al. (1999).

Furthermore, for this paper the adopted proximity weighting scheme is not used to represent context association for the LSA. The scheme is rather to generate viable context from the HTML text association within the tags or a sentence as a whole. The objective is to justify the selection of the sentences or phrases from the HTML text hence used to add to the appeal factor in the automatic image tagging.

## 4 CONCLUSIONS

The results produced thus far look potentially promising when run on 20 sampled websites. The application produced three layered descriptions for every image successfully although semantically the value of each varied. However, this preliminary indication has not been proven with the real users which should provide data that help measure the methodological significance. Furthermore, the consistency of the results seems to be affected by the weight of text proportion as well as the design quality of the page based on the Web Accessibility Initiative (WAI) standard benchmark.

An experiment to verify this approach is being undertaken in two stages. 1) Control experiments on normally sighted users, which will attempt to evaluate the generated description based on feedback from the sighted users. This will be used as the benchmark measure for the result of the second stage. 2) Acceptance experiment on the target group. This experiment will attempt to evaluate the real users experience by comparing between the pages

with and without the adaptation applied. This will provide information which may prove the significance of the hypothesis of this study.

The position of this paper is that images on web sites need to perform their intended purpose for blind users as much as for sighted users. It is therefore inappropriate to simply remove image tags as suggested by some (Shinohara & Tenenberg, 2009). It is also more important to present the purpose of images concisely and in an appealing way to the blind user. In some cases this may result in the replacement of the image with a link – when the image is merely being used as navigational cue. However, when the image is part of the page's information it is necessary to retain the image and introduce an alt tag that explains clearly and concisely the functional relationship of the image to the text. So for example for an image of a car in a used car advertisement, if the image is of the actual car it will be retained and the alt tag replaced with a phrase such as "An image of the car advertised." This then allows the blind user to know the purpose of the image and obtain help from a sighted person in determining the value of the content. However, if for example the image is a generic one it might be removed.

## REFERENCES


Bigham, J., Kaminsky R., Ladner, R., Danielsson, O., & Wempton, G. (2006) Webinsight: Making web images accessible. In *Proceedings of 8th International ACM SIGACCESS Conference on Computers and Accessibility*, 181-188.

Bigham, J. (2007) Increasing Web Accessibility by Automatically Judging Alternative Text Quality. In *Proceedings of IUI 07 Conference on Intelligent User Interfaces*, 162-169.

Caldwell, B., Cooper, M., Reid, L., & Vanderheiden, G. (2008) *Web Content Accessibility Guidelines 2.0,* Retrieved September 23 , 2009 from http://www.w3.org/TR/WCAG20/

Carson, C. & Ogle, V. E (1996) Storage and Retrieval of Feature Data for a Very Large Online Image Collection. In *IEEE Computer Society Bulletin of the Technical Committee on Data Engineering*, 19(4).

Cascia, L, Sethi, S. & Sclaroff, S. (1998) Combining textual and visual cues for content-based image retrieval on the world wide web. In *IEEE Workshop on Content-Based Access of Image and Video Libraries*.

Croft, B., Metzler, D. & Strohman, T. (2009) *Search Engines: Information Retrieval in Practice*. Addison-Wesley.

Datta, R., Li, J. & Wang, J. Z. (2005) Content-based image retrieval: approaches and trends of the new age. *In Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval*, 253–262.

Deewester, S. Dumais, S., et al. (1990) Indexing by latent semantic analysis, *Journal of Americ. Soc. Info. Sci.*, vol 41, 391-407.

Gabrilovich, E. & Markovitch, S. (2007). Computingsemantic relatedness using Wikipedia-based explicit semantic analysis. In *proceedings of IJCAI*, 1606-1611.

King, A., Evans, G., & Blenkhorn, P., (2004) WebbIE: a web browser for visually impaired people. In S. Keates, P. J. Clarkson, P. Langdon, & P. Robinson (Eds)., In *Proceedings of the 2nd Cambridge Workshop on Universal Access and Assistive Technology (pp. 35 – 44*), Springer-Verlag, London,.

Li, J. & Wang, J. Z. (2008). Real-time computerized annotation of pictures. *IEEE Trans. Pattern Anal. Mach. Intell*. 30(6), 985–1002.

Macdonald, C. & Ounis, I. (2010) Global Statistics in Proximity Weighting Models. In *Proceedings of Web N-gram Workshop at SIGIR 2010*, Geneva.

Nakov, P., Popova, A. & Mateev, P. (2001) Weight functions impact on LSA performance, In *Proceedings of the Euro Conference Recent Advances in Natural Language Processing* (RANLP), 187–193.

Paek, S. & Smith, J.R.(1998) Detecting Image Purpose. In World-Wide Web Documents. In *Proceedings of IS&T/SPIE Symposium on Electronic Imaging*: *Science and Technology – Document Recognition.* San Jose: CA.

Petrie, H., Harrison, C. & Dev. S. (2005) Describing images on the web: a survey of current practice and prospects for the future. In *Proceedings of Computer Human Interaction International (CHI)*.

Petrie, H., O'Neill, A-M. and Colwell, C. (2002). Computer access by visually impaired people. In A.Kent and J.G. Williams (Eds.), *Encyclopedia of Microcomputers*, vol 28. New York: Marcel Dekker.

Petrie, H. and Kheir, O. (2007) The relationship between accessibility and usability of websites. In *Proceedings of Computer Human Interaction International (CHI)*

Russell, B., Torralba, A., Murphy, K., & Freeman, W., (2005) LabelMe: a database and web-based tool for image annotation (*Technical Report*). Cambridge, MA: MIT Press.

Sclaroff, S., Cascia, L. & Sethi, S. (1999) Using Textual and Visual Cues for Content-Based Image Retrieval from the World Wide Web, *Image Understanding*, 75(2), 86-98.

Shinohara, K., & Tenenberg, J. (2007). Observing sara: a case study of a blind person's interactions with technology. *In Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility*. New York: ACM Press. 171–178

Shinohara, K. & Tenenberg, J., (2009). A blind person's interactions with technology, *Communications of the ACM*, 58-66.

Stork, D. G. (1999) The Open Mind Initiative. *IEEE Intelligent Systems & Their Applications*, 14(3), 19-20.

Stork, D. G. & Lam C. P. (2000) Open Mind Animals: *Ensuring the quality of data openly contributed over the World Wide Web. AAAI Workshop on Learning with Imbalanced Data Sets*, pp 4-9.

Sutton, R. I & Staw, B. M. (1995) What theory is not *Administrative Science Quaterly*, vol 40, 371-384.

Tao, T. & Zhai, C. (2007) An exploration of proximity measures information retrieval. In *Proceedings of SIGIR' 2007*, 295-302.

Von Ahn, L. & Dabbish, L. (2004). Labeling images with a computer game. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*.