

Automatic Image Annotation using Image clustering in Multi – Agent Society

Abbas Akram Khorsheed
University of Mustansiriyah,
College of science, Computer
Science Department, Iraq

Hussein Chible
Doctoral School of science and
Technology, Lebanese
University, Beirut, Lebanon

Giorgio Giacinto
Cagliari Univesity,
Electronic Engineering
Department, Italy

Abstract: Image annotation is the natural language description for essential blobs within an Image. Many methodologies have recruited to reveal image semantics and represent it as annotation. Due to the exclusionary growth of number of images distributed over massive repositories, the task of manual annotation is tedious and over killing in term of times and efforts. This paper is presenting automatic annotation system based on the interaction between intelligent agents. Agent interaction is synonym to socialization behavior dominating Agent society. The presented system is exploiting knowledge evolution revenue due to the socialization to charge up the annotation process

Keywords: Social effect, Automatic annotation, intelligent agents

1. INTRODUCTION

Type Style and Fonts Image annotation is one of the most used methods to retrieve images from enterprise repositories by matching user text queries to these annotations. In general, annotation is represented as metadata or keywords assigned to digital images based on image contents [1][2].

More than 200 billion images are accessible online and the number is continuously growing [3] due to the numerous number of sources as digital cameras, mobile phones and other devices. This brings a great challenge in retrieving designated images which are identified by a unique number over the internet, i.e., the URI (Unique Resource Identifier), that is used to access each image over the web. Social decision theory [1] extends the theory of individual decisions to decisions made by the interaction of a group of agents.

Recent systems like Lable Me and Amazon mechanical turk distribute image annotation and evaluation tasks to Internet users. The volume of annotations generated from such crowd-sourcing techniques helps reduce the burden on experts without significantly sacrificing the quality of annotations. The annotators are provided with detailed instructions on how to best select labels that can be directly used for concept modeling. This ensures that relatively good quality annotations are generated for object detection, and relevance estimation tasks. It is shown that crowd-sourcing is a reasonable substitute for repetitive expert annotations, when there is high agreement among annotators.

Other sources of image annotations are collaborative games and social media sharing which undoubtedly represent the fastest growing labeled image collections in the world[3][4].

In this paper the multi Agent paradigm is proposed to simulate the social behavior of humans in developing knowledge regarding certain subjects. The JADE (Java Agent Development) environment has been used because it allows building multi agent platforms thanks to the utilities and wide spectrum classes provided by that environment.

2. RELATED WORKS

Many models were produced to characterize automatic image annotation frameworks. The exact specification of automatic image annotation is not yet established and researchers are investigating different approaches continually. Anyway, recent approaches can be categorized into two categories: one category is the semantic interpretation of image contents, while the other category is drawn from the epistemology field, where knowledge is revealed from the interaction among sources of knowledge. The society can be represented as a network of knowledge resources, and knowledge can be sustained or rejected upon the interaction among these resources. In this section we will focus upon previous efforts within the second category due to the orientation of this paper.

In [5] a model has been proposed to formulize the growth dynamics in social networks; in this model a great attention has been presented to the effect of node behavior, and how it affects the behavior of other nodes, and this eventually will affect the growth of the network. In term of knowledge evolution due to socialization; this model has a lot in common with our approach, though it has nothing to do with image retrieval system. The key similarities are:

1. The behavior-awareness where the interaction of node (i.e., the co-author s) with certain events (e.g., papers) is to be realized as a potential relationship among those nodes. In fact this approach develops knowledge at the network level, which helps increasing the growth factor of social network and, eventually, the productivity of such a network.
2. The clustering-coefficients where the tendency of grouping is related to the factor compose of these coefficients.

In [6] an ontological approach was presented to accomplish a computing model aimed to annotate images on two levels: Image Annotation and Annotation of Annotations; this model.

is focused on queries for annotations using the National Cancer Institute’s Cancer Biomedical Informatics Grid’s (caBIG) Annotation and Image Markup (AIM) project

The AIM project defines an ontology of annotations and image markup, a UML information model and provides the extensible markup language (XML) artifacts for creating them. A long-term vision of the AIM project is for large collections of annotations to be created in conjunction with the already large collections of clinical and research medical images. This will allow query of annotation, not only for the retrieval of relevant images, but also for the correlation of image observations and their characteristics with biomedical data including genomic expression.

In that paper many concepts are coherent with what we presented in our work in the area of retrieving images based on associated annotations, but this approach does not introduce autonomous annotation in any context, and it does not consider the behavior of image requesters; knowledge can’t be developed to cluster images which is a crucial element in automatic image annotation strategy.

The model presented in [6] exploits annotations to build a semantic network among images, while our work provides autonomous annotation schema based on the behavioral interpretation of the user. The AIM project can be integrated with what we are presenting to provide consistent ontological environment for image retrieval and annotations. The same annotation context is presented by [7] and [4] but both depend on the retrieval and extraction of knowledge from the resources available on the global net.

In [8] a novel system is presented to exploit the format of multimedia sharing web sites in order to discover the underlying structure; this has been used to allow later, more sophisticated mining tasks for these sites to infer knowledge about certain images. Again, we have many features in common with these approaches, but still the effect of the behavioral responses of the users is absent.

In [9] a study for establishing a stable architecture for socialization is conducted and conclusion has been reached out along this study which is: in a society of agents there are three main parameters that enforce the stabilization of the architecture; these are: take on roles, play roles and locate in some society organization at all time. In our proposal, the society composed by agents is maintained stable by strict discipline through which roles are fairly distributed, and all agents are capable of playing these roles by accurate interpretation of client behavior. Furthermore, we adopt fixed organizational distribution of the agents which sustain the stability. In our proposal, the specification of the problem domain has different characterization due to the potential tendency toward clusterization on two different levels: the host level, and the network level. This approach has its roots back to [10] where a study addressed the fault assumption of regarding multi-agent systems as single learning system which is a wrong assumption due to the intuitive tendency to introduce social activity with neighbors rather than communicating with other far agents. This dual capabilities of an agent’s referencing, i.e., self-referential, and social-referential, has been presented by [10] as a bi-referential model, in which each referencing capability is implemented by an evolutionary computation method of classifier system.

In our referential model the evaluation function is global and updated on the fly by delivering knowledge to central a repository that holds the annotation for images. The

annotations are revealed and referenced based on a confidence degree assigned to that annotation. In our referential model, the behavior of the evaluation function is dynamic due to the continuous change of confidence degree of annotation; this is due to the activities produced by the client clusterization behavior (i.e., the self-referential model).

Interactive query for images’ content by semantic descriptors is an effort presented in [11]; this effort introduced a distributed content-based image query system (DCBIQ) based on the WWW. A model was proposed to integrate knowledge from image processing, semantic descriptor, multi-agent, and WWW navigation. Again in this model the image content plays the essential role in describing the image, thus low level extraction methodologies are more important than the opinion of the social communities which are using it.

In our proposal, the knowledge obtained by social interaction is more important than low level features like colors, textures or spatial relationships, and even semantic interpretation of image contents is not important as the social opinion about the image and its relation to other images or domains.

In [12] an attractive model is presented where a web-based image digital library is proposed; in this library agent system was used to traverse part of the web page looking for images that fit certain criteria. The methodology used by the agent is based on detecting URLs within web pages that refer to images, and when such URLs are encountered, then the text that is associated to that image is inferred for correlation with other features such as topic name, domain that this image falls in, or any other matching criteria. In our proposal the same ontology for allocating text accompanied the image is used as the following matched methodologies:

- 1-
$$\forall \text{Image} \exists \text{tag} \exists \text{txt} ((\text{presenting}(\text{image}, \text{tag}) \text{ AND } \text{Asso}(\text{tag}, \text{txt})) \longrightarrow \text{Asso}(\text{image}, \text{txt}) \text{ AND } \text{Select}(\text{txt}))$$
- 2-
$$\exists \text{paragraph} \exists \text{hyper} ((\text{hyper} \in \text{paragraph}) \longrightarrow \text{Select}(\text{Paragraph}))$$
- 3-
$$\exists \text{page} \exists \text{time} \exists \text{!title} ((\text{image} \in \text{page}) \text{ AND } \text{has}(\text{title}, \text{page}) \longrightarrow \text{Associate}(\text{image}, \text{title}) \text{ AND } \text{Select}(\text{title}))$$

The main novelty of our approach is that we don’t design a mining agent that is responsible on inferring web pages, but we exploit Google search APIs which are published over the web. The only web page we analyze is the results of the Google search APIs and don’t investigate individual pages.

3. THE PROPOSED SYSTEM

This paper will focus on new category which is the dominant tags of the image as it is recognized by the society. Image repository (RTI) is a database holding labeled images (i.e. images tagged with annotation), thus it can be represented by eq.1

$$R_T^I = \sum_{i=1}^k (I_i, Anno_i) \quad \text{eq.1}$$

Where $Anno_i = \sum_1^m c_j$ collection of concepts (i.e., these concepts are keywords, tagged or labels), and let the query made by the user in order to request images is representing by the following equation:

$$Query_i = \sum_1^n q_j \quad \text{eq.2}$$

$\forall \text{ indexed}([image]_i) \exists c \in Anno([image]_i)$ so that $c \notin Anno([image]_j)$ for all $j \neq i$

$\text{SearchResult}(Query_i) = \sum_{j=1}^M images_j$ iff

$$Query_i \cap Anno(image_j) \neq \emptyset$$

and $OSF(Query_i) \neq c$

OR

$\text{SearchResult}(Query_i) = image_i$ iff

$$OSF(Query) = c$$

Behavior	Weight
Saved	5
Selected, saved	4
Revisited, saved	3
Revisited	2
Highlighted	1

User interaction behavior with the resultant list of images is weighted according to above table.

Hypothesis 1: Automatic annotation member is broadcasting candidate annotation for queried images.

Prove:

Let $f(query)$ be a mapping function that maps images from the huge repositories spread over the internet to the desired image list requested by the user, such that

$$f: Re \longrightarrow De$$

Where Re is the huge repository over the internet and De is the desire domain where resultant of $f(query)$ satisfies client request. The input to this function is the query entered by requestor and the output is a scalar value represents confidence degree

$f(query) = 0$ Where 0: not desired
 and 1: desired, thus

$$0 = De \quad \left\{ \begin{array}{l} v < a < 1 \\ 0 \neq De \end{array} \right.$$

Where

$v = T.W_{Agent}^{image}$ Which is the total weight produced by the interaction between the requestor client and the resultant list of images. If $v \geq \text{threshold}$ then agent will broadcast a data structure composed of the following fields (Image URI , query, $f(query)$).

Definition: dominant annotation is the candidate new annotation for image being queried by society of agents, where $f(query)$ for 1 want here sigmoid function to be the decision function to decide that certain annotation is to be added to the image annotation list .

Hypothesis 2: social group add new annotations to image

Let :

$$Query_A \cap Anno(image_j) = S_A$$

$$\text{and } = \sum_{k=1}^K \text{weight}(\text{behavior})$$

$$Query_B \cap Anno(image_j) = S_B$$

$$\text{and } = \sum_{i=1}^L \text{weight}(\text{behavior})$$

⋮

$$Query_n \cap Anno(image_j) = S_n$$

$$\text{and } T.W_{Agent_n}^{image} = \sum_{m=1}^M \text{weight}(\text{behavior})$$

$T.W_{Agent_{ij}}^{image_j}$ is Total weight produced by $Agent_{ij}$ for $image_j$

Then

Added annotation set $S = S_A \cup S_B \dots \cup S_n$ is a set of new valuable annotations to be added

to $image_j$ with a binding value $T.W_{Agent_{ij}}^{image_j}$, hence the resultant set is only a candidate

annotation, it has to be dominant to get corresponding image get indexed with.

4. SOCIAL EFFECT

Social effect over Automatic Annotation Society will be treated in this section. In social environment, members are investigating propositions based on total weight granted by trusted members of the society. In this proposal we assume that all members are trustful and other members of the society are considering their weight evenly. Let us first define a new function that describes the acceptance of the society for the candidate annotation to be a dominant annotation, and the corresponding image can be indexed with. In this paper we propose the social effectiveness function to be a sigmoid function, due to the properties of this function especially the continuity and flexibility, hence $f(query)$ is defined as

$$SocialAcc(Image) = \frac{1}{1 + e^{-Su(query)}}$$

Where $Su(query)$ is a total weight gained from all agents involved in the automatic annotation system and it is represented as the following:

$$Su(query) = v_1 + v_2 + \dots + v_n$$

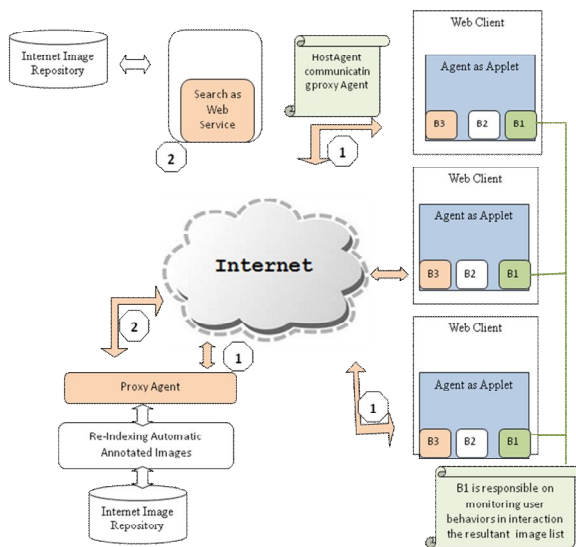


Figure 1: General scheme of social basic automatic annotation system

The role of Agent in this proposal is characterized by three behaviors, as presented in figure(2) and are briefed as the following:

Web Service Integration (B3): this behavior encapsulates web service integration functionalities by implementing SOAP based invocation to Google web service.

Socialization Behavior(B2): this behavior is responsible on socializing other agents within the platform to determine dominant tags for an image.

User behavior monitoring (B1): this module is responsible for monitoring selections made by the user after querying the Google search Web service. Highlighted images are grouped in clusters and socialization behavior is signaled. The

following behaviors are considered: Highlighted images, selected images, revisited images and saved images.

Proxy Agent: this agent is responsible of initiating the communication session over the internet. Proxy Agent is a crucial element in grant multi-agent system the ability to communicate over the internet. This Agent resides at the server side.

Host Agent: this Agent is an instance constructed at the client side and monitor his/her behavior and report back to Proxy Agent his observations.

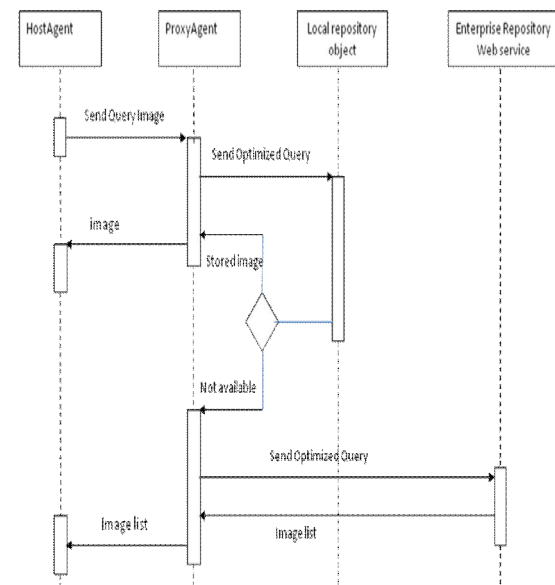


Figure 2: sequence diagram of complete 2-tier image annotation session

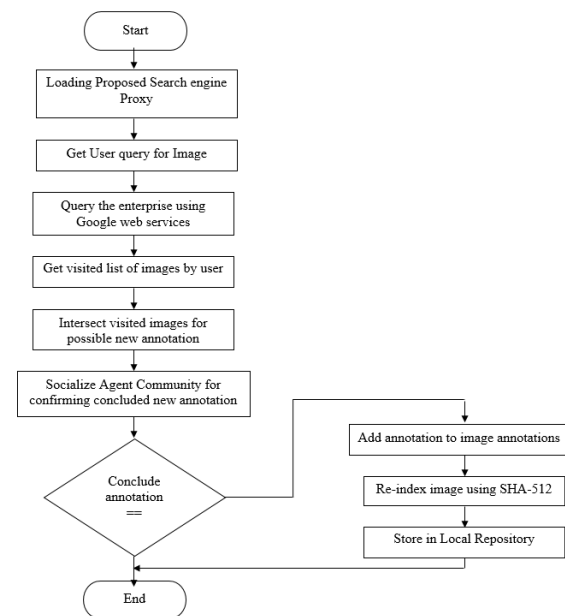


Figure 3: Automatic Image Annotation Procedure Using Multi-Agent socialization

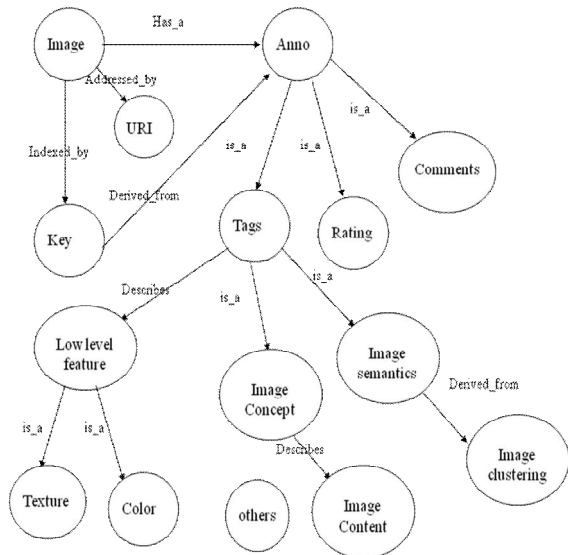


Figure 4: Ontology to conceptualize Image Annotation

5. RESULTS

Example:

Phase 1:

client₁ has posted the following query through the Chrome internet explorer

query₁ = { car, rental, company, race }

After posting that query, 37,212 images have been listed in the internet explorer.

client₁ has selected and saved the following image



The session manager agent, which has HTTP listener, captured the URL or URI corresponding that image.

The following is the URL

Imgurl: http://www.koopman-racing.nl/images/sd2_1559.jpg

This image will be indexed using **query₁**.

www.ijcat.com

Phase 2:

client₂ and *client₃* have posted queries as the following

query₂ = { aspiring, driver, koopman, race }
with max weight (5)

query₃ = { race, rental, koopman, car }
with weight (4)

query₁ ∪ query₂ = { aspiring, driver, koopman }

these annotations are to be more convenient to be used as indexing due to its weight factor and by socializing it to other clients like *client₃* the highest effective annotations will be **{ aspiring, driver }**, This is for the same URL.

<https://secure.booking.com/confirmation.fr.htm?aid=350433;label=edr-xmlvswl-fr-users;sid=05686c51355c9e5ba1e2a8d843e2c461;dcid=2;bn=608419725;hostname=hotels.edreams.fr;pincode=6604#print>

Algorithm1:

Automate Image Search Using Google

- Procedure: Search Web
- Input: query As String
- Output: array of imgurls
- Begin
- Initialize user Query = query;
- Initialize GoSearchConnection as URLConnection to Google URL + user Query;
- Set GoSearchConnection Properties as
 - Method = 'GET';
 - Char-set = 'utf-8';
 - User-Agent = 'Mozilla-4.0';
- GoSearchConnection. Open;
- Get input Stream from GoSearchConnection to stream Reader;
- while stream Reader has imgurl do
 - add current imgurl to imgurl_list;
- return imgurl_list;

End;



Query: sport car	
index	Imgurl
1	http://investopedia.com/articles/financialedge/Rangerover.jpg
2	http://www.bestautowallpaper.com/wp-content/uploads/2012/03/Mazda_Furai_Race_Car_Concept_Best_Auto_Wallpaper_HD_by_engineslover.jpg
3	http://www.pdd.co.uk/media/114073/denza_car_design.jpg
4	http://wallpapersget.com/wallpapers/2012/03/car-bmw-328-hommage-wallpaper-1080x1920.jpg
5	http://img.gwkerassets.com/img/186u39dbyn5b3jpg/original.jpg
6	http://fp.images.autos.msn.com/Media/425x255/be/be10a5c6a0e5443a8217388ab76eb41a.jpg

Query: black fancy race car	
index	Imgurl
1	http://www.carspictureswallpaper.com/wp-content/uploads/images/33506-desktop-use-only-tags-supercar-sport-car-car-orange-black-luxury-car_531x331.jpg
2	http://www.autotrader.co.uk/EDITORIAL/editorial_images/celeb_cars_montage.jpg
3	http://popsop.com/wp-content/uploads/bmw_mark_lewis_francis_02.jpg
4	http://db2.stb.s-msn.com/f/1C/87789A1A156480771D5CECA36E37B.jpg
5	http://wallpapersus.com/wp-content/uploads/2012/11/BMW-White-Racing-Car.jpg
6	http://fp.images.autos.msn.com/Media/425x255/be/be10a5c6a0e5443a8217388ab76eb41a.jpg

Query: celebrity car sport	
index	Imgurl
1	http://media.caranddriver.com/images/12q4/477958/2014-chevrolet-corvette-c7-secret-info-and-renderings-news-car-and-driver-photo-493641-c-450x274-1.jpg
2	http://www.bestautowallpaper.com/wp-content/uploads/2012/03/Mazda_Furai_Race_Car_Concept_Best_Auto_Wallpaper_HD_by_engineslover.jpg
3	http://www.pdd.co.uk/media/114073/denza_car_design.jpg
4	http://wallpapersget.com/wallpapers/2012/03/car-bmw-328-hommage-wallpaper-1080x1920.jpg
5	http://images.thetruthaboutcars.com/2010/02/Ferrari458-tuned.jpg
6	http://www.wpdaddy.com/previews/luxury-car-theme_big.jpg

Algorithm 2:

Reveal Local Knowledge

- Input: Selected imageUrl_list
- Output: weighted imageUrl_list
- Begin
- For each image in image Url list Do
- begin
- Capture mouse and keyboard events
- Assign weight to image
- end
- End.

Algorithm 3

: Broadcast local knowledge

- Input: weighted imageUrl_list
- Begin
- Instantiate msg from ACLMessage;
- Set msg.receiver to be address of global Agent
- Set msg.content to be weighted imageUrl_list and the Query;
- Send msg;
- End

Algorithm 4:

Intersect broadcast knowledge

- Input: msges[] as Array of Agents' Messages
- Begin
- Initial CommonVisited List as String Array
- Initial SumWeight as integer Array
- For all messages in msges[]
- tempMsg = nextMsg in msges[]
- for all messages in msges[] and NOT tempMsg do
- find shared imageUrl and add it to CommonVisitedList.
- Sum total weight and add it to SumWeight in index manner.
- End.

After intersecting queries from different Agents, the following URI

- <http://wallpapersget.com/wallpapers/2012/03/car-bmw-328-hommage-wallpaper-1080x1920.jpg>

will be indexed using key {race, fancy, celebrity and sport .

6. CONCLUSIONS

After investigating a bunch of papers published within the same topic of our proposal, we found correlation in basic terminologies, but with distinct methodologies. Many models have been introduced to develop knowledge about retrieved images like what we introduced here and the significant features of Agent-based system are also exploited but the key differences between all these efforts and what we devised in our proposal can be summarized:

1-A Multi-Agent system has been deployed on two levels: host level and network level to develop knowledge regarding certain images, other approaches target mainly behavioral aspects of network interactions rather than host based.

2-Annotation is generated autonomously and a confidence value is assigned to each annotation; this value represents the acceptance of society for this annotation as a key index for associated image.

3-Third party web based tools has been included (i.e., the Google search engine APIs) while all other approaches tend to design custom search software modules. Google has a very massive repository of images, thus it is more convenient to address this repository rather than inferring other repositories or web pages. Furthermore, Google search engine receives millions of request for images in multiple subjects, thus this will assist, statistically, revealing more reliable annotations.

4-An image is annotated, in our proposal, not on the basis of the graphical objects in the image or the low level features, but on the basis of its relation to the environment, for example an image could have some planets and this image can be interpreted using low level features and semantic contents as to relate to planet science, flowers, garden or some of the like.

In our approach the planets image can be categorized into drug, medicine, health or so on; this is due to society opinion.

Other approaches Index this image based on its low level feature and its composed visual objects.

7. ACKNOWLEDGMENTS

My thanks to the University of Mustansiriyah – Iraq, that gives me the possibility to support and continue my PhD.

8. REFERENCES

- [1] NursuriatiJamil and SitiAisyahSa'adan, Proceeding, "Visual Informatics: Bridging Research and Practice", visual informatics conference, IVIC 2009, Kuala Lumpur, Malaysia, 2009.
- [2] Masashi Inoue, "On the Need for Annotation-Based Image Retrieval", National Institute of Informatics, Tokyo, Japan, 2006
- [3] Aixin Sun, Sourav S. Bhowmick, Khanh Tran Nam Nguyen, "Tag-Based Social Image Retrieval: An Empirical Evaluation", published in American Society for Information Science and Technology, USA, 2011.
- [4] NeelaSawant, Jia Li and James Z. Wang, "Automatic Image Semantic Interpretation Using Social Action and Tagging Data", survey, Pennsylvania State University, USA,2011.
- [5] Baojun Qiu, Kristinka Ivanova, John Yen, and Peng Liu, "Behavior Evolution and Event-driven Growth Dynamics in Social Networks"
- [6] David S. Channin MD, Pattanasak Mongkolwat, Vladimir Kleper, and Daniel L. Rubin, "Computing Human Image Annotation"
- [7] Chen-Ue Lee, Von-Wun Soo, and Yi-Ting Fu, "How to Annotate an Image? The Need of an Image Annotation Guide Agent"
- [8] Nikolaos Papadakis, Klimis Ntalianis, Anastasios Doulamis, and George Stamoulis, "An Automatic Multi-Agent Web Image and Associated Keywords Retrieval System"
- [9] Jian Zhang, Guang-Zhou zeng, and Zhi-Feng Li, " The Study of a Sociality Agent Architecture Based On Role"
- [10] Takayuki Shiose, Tetsuo Sawaragi, Osamu Katai, and Michio Okada, "Dynamics of Reciprocal Learning by Bi-Referential Model within Multiagent Systems"
- [11] Weng Zu Mao, and David A. Bell, "Interactive Query for Image Contents by Semantic Descriptors and Multi-Agent".
- [12] Jesus Favela and Victoria Meza, "Image-retrieval agent: integrating image content and text".