Personal Image Tagging: a Game-based Approach

Jakub Šimko
Institute of Informatics and Software Engineering,
Faculty of Informatics and Information Technologies,
Slovak University of Technology,
Bratislava, Slovak Republic
jsimko@fiit.stuba.sk

Mária Bieliková
Institute of Informatics and Software Engineering,
Faculty of Informatics and Information Technologies,
Slovak University of Technology,
Bratislava, Slovak Republic
bielik@fiit.stuba.sk

ABSTRACT

An effective search and organization of personal multimedia repositories demands very specific, owner-related metadata (e.g. person names, places, events). Only the resource owners and their social circles are able to provide these metadata, but are often not motivated to do so. To increase their motivation, we introduce a game-based personal image annotation framework. Today's crowd-based games with a purpose (GWAP) are able to harness human labor to acquire metadata for general domain multimedia resources. However, their deployment in such specific domain as one's personal multimedia is hindered by the issue of extremely small number of competent potential workers. Therefore, a traditional cross-worker agreement filtering of wrong answers cannot be effectively applied. In our approach (which is an extension of an existing image tagging game), we overcome this issue by inviting user to play for his own good i.e. annotate his own images while enjoying the game. This additional motivation causes an overall increase of quality of user-generated metadata and thus allows the use of less strict tag extraction algorithms for producing final set of tags. We show that our approach is able to yield valid image annotations, specific to the context of the resource owner. We also examine its performance with different types of annotated images and tags.

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: Content Analysis and Indexing; K.8 [**Personal Computing**]: Games

General Terms

Design, Experimentation

Keywords

human computing, games with a purpose, multimedia metadata, personal multimedia

I-SEMANTICS 2012, 8th Int. Conf. on Semantic Systems, Sept. 5-7, 2012, Graz, Austria.

1. INTRODUCTION AND MOTIVATION

In this work, we present a human-oriented, game-based approach to acquisition of textual metadata for *personal image archives*. Personal multimedia are characterized by their limited public availability and specificity. However, just as their general counterparts, they need to be properly decorated with metadata. This enables effective search, browsing or even automated storytelling [18] for their owners and social circles.

Acquisition of image metadata over general domain (general concepts) is problematic, but for personal images, it gets even worse. While automated approaches to image annotation can identify specific visual features (such as human faces) or categorize images, they are unable to extract the information specific to its owner such as person's name, place or event, when the picture was taken.

Human-performed annotation generally provides better quality metadata than automated means. The issue of (non-monetary) motivation of people to provide annotations has been addressed by crowdsourcing approaches (e.g. games with a purpose – GWAP). However, existing approaches are short handed when it comes to personal multimedia archives. This is mainly due to privacy reasons and also the fact that only narrow group of people is usually able to provide sufficient specific information needed for this type of content. Only the social-based approaches (e.g. tagging or commenting images of one's social circle) are able to provide metadata addressing image owner's context. However, these are very sparse, irregular and often noisy due to sentiment, humor, etc., contained in social comments from which they are extracted.

A main contribution of this work is a GWAP-based personal image tagging approach that produces image tags that are *valid* and also *specific* for the user (and therefore usable in personal image repository management). It overcomes the problems described above with a combination of three kinds of player motivation to participate in the process:

1. The player enjoys interaction with a content he likes (i.e. one's personal images or images of one's social group). The game is considered also as an image presentation tool.
2. The player is challenged by the game.
3. The player gets his images annotated.

The idea is to let the user to play an image annotation game, but with his own images and for his sake (his metadata,

which can be used e.g. for better image search). The game also aims to exploit a possible scenario when a player is sharing the game with his social circle (to share the images and acquire more metadata for his use).

Our approach is an extension of our existing game with a purpose – the *PexAce* (a modification of a popular board game – the *Concentration*), where players collect identical card pairs and help themselves by annotating the images on cards. The PexAce game has already proved its usefulness in general image domain [12]. In this work however, the game functions over player's own images and extracts tag for them. The original multiple-player-agreement tag filtering approach (used traditionally in games with a purpose) has been modified and extended by other heuristics that take the specific nature of personal images into consideration: existing organization of images to albums, repetition of specific person names, places and other concepts in the annotations created during the gameplay.

We have performed a combined quantitative-qualitative study to evaluate our approach. We show that the game is able to provide *valid* and also *image-owner-specific* metadata to images used in the game, while retaining a key feature of *low number of participants needed to play.* We further examine the types of tags extracted as well as performance of the game over different types of images. Also, with respect to the possible scenario of sharing the game among social circles, we measured how does the game perform, when players are unaware of its purpose (but still play over "their" images).

2. RELATED WORK

Needs and preferences of users in terms of dealing with their personal images have been well described by a qualitative study of Vainio et al. [13]. When asked, users admitted that they like to interact with their content (watching, but also editing), but not with its metadata. They realize the value of metadata, especially with larger repositories, but are generally not willing to systematically create them. The most common practice in organizing a personal image repository is just labeling the whole collections (albums) with names or sorting images chronologically. Therefore, users must rely on time-consuming sequential browsing when trying to find a particular image. An important closure was, that users would primary welcome metadata about persons in the images, person who took the pictures and situational context of pictures (e.g. information about places and events) [13]. The latter strongly implied the directions of our research: in our study, we examine the performance of our approach with these types of metadata.

Image metadata can be acquired either by automatic or human-oriented means. Despite their volatile nature in terms of quality, automated metadata acquisition approaches are used for annotation of large image collections of the general domain and are essential for the web-scale image search. Use of automated means in the personal image metadata acquisition scenario may be possible, but only in the supporting role, as they are able to work only with general concepts. They can carry out a general category identification for images, while leaving the more specific metadata identification to humans.

The most common practice for image web search is to analyze the context of images [6, 8]. Personal images might have a textual context (e.g. they are being shared via social network, got commented), but usually they do not. More promising for this type of multimedia would be the analysis of visual features of the images, for which several approaches have been designed. Employing image segmentation and large scale vocabularies, Duygulu and Barnard's experiments yielded 70% accuracy in associated tag prediction over Corel 5K dataset [4], which was further improved after introducing a probabilistic models [11]. Joint use of global and local feature identification in the images also helped the prediction [1, 5]. An effective narrow categorization was also achieved using SVM [3] or Bayes point machine [2].

Together, all automated approaches also require very wide training sets to perform even moderately in broad domain range. Within the personal image annotation scenario with the need for specific metadata, the acquisition of proper training sets is virtually impossible, as those require already annotated images. Nevertheless, it is imaginable that metadata acquired through human-oriented approaches (such as one described in this paper) could be used in some machine-learning-based, personal imagery scenarios (e.g. identification of persons over previously identified face regions).

Human-oriented approaches to image tagging perform qualitatively better than the automated ones. Apart from expert tagging (e.g. press image databases) or paid crowd-based services (e.g. Mechanical Turk), the approaches for user-generated image tag acquisition could be split into social tagging approaches [17] and games with a purpose (GWAP)[14]. The social-based approaches are capable of delivering specific personal metadata to images, but irregularly and with a lot of noise. However, social tagging also implies a notion which we took in our research: an *eventual human based approach to personal metadata acquisition should source it's contributors through socialization phenomena.*

The games with a purpose (as coined by Luis von Ahn [14]) represent a family of crowd-based approaches utilized for various types of HITs (human intelligence tasks) including the multimedia metadata acquisition. The first and well known ESP Game, created by von Ahn, employed two players that had to come up with the same word describing an image (or) given to them, thus validated each other [14]. Several other similar games followed this design, aiming to improve the quality of tags [9] or examine the properties of acquired tags [7].

One of the major advantages of GWAPs against the social tagging is that they provide more control over assignment of HITs to users. They can thus allocate the available manpower to annotation of resources that has been lacking the metadata.

Typically, GWAPs use the redundant task solving (as in the general crowdsourcing) to ensure their output quality [16]. However, this requires a certain number of players able to provide metadata to a certain resource. This imposes a problem which must be overcame prior to their eventual use in personal images scenario: the insufficient number of players able to do solve the task (provide enough specific metadata).

One of the features characterizing GWAPs is to what extent they encapsulate the purpose of the game, i.e. how well the purpose is hidden from the player. Markus Krause et. al. argues for the highest possible degree of encapsulation and demonstrates it on the case of the OnToGalaxy game (an ontology population game which is "concealed" as a space shooter action game). Krause notes that an appar-

ent purpose of a game potentially discourages players from playing because they will more likely consider the game as a work which will ultimately lead to game's misfortune [10]. However, in our work, we aimed for the exact opposite, i.e. that the players must be told about the purpose of the game in order to improve its useful outcome, which can directly serve for particular user or his friend or relative.

3. GAME-BASED IMAGE ANNOTATION

To enable more effective maintenance of personal multimedia repositories, we devised a method for personal image metadata acquisition, which can be split into two parts:

- *The game* as a means of image annotation acquisition.
- *The tag extraction procedures* working over collected annotations created in the game.

The gameplay system and the context of the use of the game in our extension *implies the high potential quality of these annotations* due to three types of user motivation:

1. Players interact with a familiar content – personal images. Therefore they know many specific information about them and also their context. It also increases the joy of gameplay itself.
2. Players are challenged by the game. They wish to reach a highest possible score to overdo themselves or members of their social circle.
3. Player who knows about the true purpose of the game – annotating their own image set – provide annotations not just for sake of the game, but also with respect to the future use of tags in image search and organization.

3.1 The Image Annotation Acquisition

We base our method on an existing game – the PexAce [12, 15], which we devised for image annotation in general domain. The principle of the game itself is further based on the popular memory turn-based board game called *Concentration*. In it, players seek identical images by flipping cards pairs facing down the board. If player finds an identical pair, he collects it, receives point and draws another card, otherwise he passes the turn. The cards are usually laid down in a grid with a total number of 50 to 100, so players do usually many blind attempts at the start of the game to learn the positions of images and remember them as precisely as they can.

In the PexAce, as well as in our approach, this principle is turned into *single player* and introduced the *card annotation option* for the player. Instead of having to remember, where exactly each particular image is hiding, player have the option to write down an annotation on the image when he sees it and then "stick it" to the underside of the card. Player can see this annotation at any time during the game to aid his memory.

The player gets score based on his own performance: the lower number of flips he needs to finish the game, the more points he receives. He may also increase his score by finishing the game earlier. His score is afterward used as a reward for competing with other players or within player's own self-challenge. In original PexAce, players competed in a central ladder. However, in our approach we chose to pursue competition only among one's social network (not all players of the game) as we expect less motivation for cheating, which was experienced during the PexAce releases.

An interesting score-related feature of the game is that players may also compete with each other *even if they play with completely different sets of images*. Players cannot make the game easier by picking some obscure set of images for play (in fact, it might make them the game harder). They will always have to use either their memory or make precise annotations. This way, we can eventually pursue competition even in the centralized score ladder, while retaining the option for each player to annotate his own images.

We have implemented the game as a desktop application to easily reach local personal image repositories. Its player interface can be seen in the Figure 1. It is based on the original PexAce look: player may flip two images at a time and annotate them using the fields in the center. Later, he can show these annotations by hovering over undisclosed cards in the card grid. The image set is selected from the file system prior to the game.

The option to annotate the images implicitly *motivates players to provide annotations really characterizing the content of the image* (otherwise it would be invaluable for them). For further automated tag extraction, this is an important feature, as it significantly lowers the redundancy validation requirements of the process. Because the players interact with the *personal* content or a content to which they are socially bound, they are expected to use also specific descriptions such as person names, places, events and other contextual information, indirectly describing the image.

Other games with a purpose dealing with image annotation could also be considered to be extended to work over personal image archives. The PexAce, however, has a significant advantage of being single-player and having its scoring function independent of the quality of player annotations: the scoring independence reduces the possibilities that a player would develop a certain winning strategy in the game which could potentially hamper the purpose of the game.

3.2 The Tag Extraction

The result of the game phase of our approach is a set of raw textual annotations from each one is assigned to a certain image (in a certain image set and album) by a certain player. Our method processes the raw game annotations by automated means to extract tags. Firstly basic natural language processing methods – tokenization, lemmatization and stopwords removal – are used to create *tag suggestions*. A tag suggestion is a hypothetical "vote" of a player for a certain term to be connected with an image. It is based on an assumption that terms used in the annotation text (which could be a whole sentence or a sentence fragment) are related to the image also individually. The follow-up processing then tries to assume, how strong that relatedness is.

We have devised several heuristics that estimate if the individual terms should be assigned to the images. Given that a *tag assumption* is a quartet of *player*, *term*, *image* and *album* identifiers, an image gets decorated by a certain tag:

- If certain number (in our experiments we used 2) of players agree on the same suggestion for same image. This was the validation method used also in the PexAce. This traditional crowdsourcing heuristics is very restrictive. It relies on sufficiently large number of

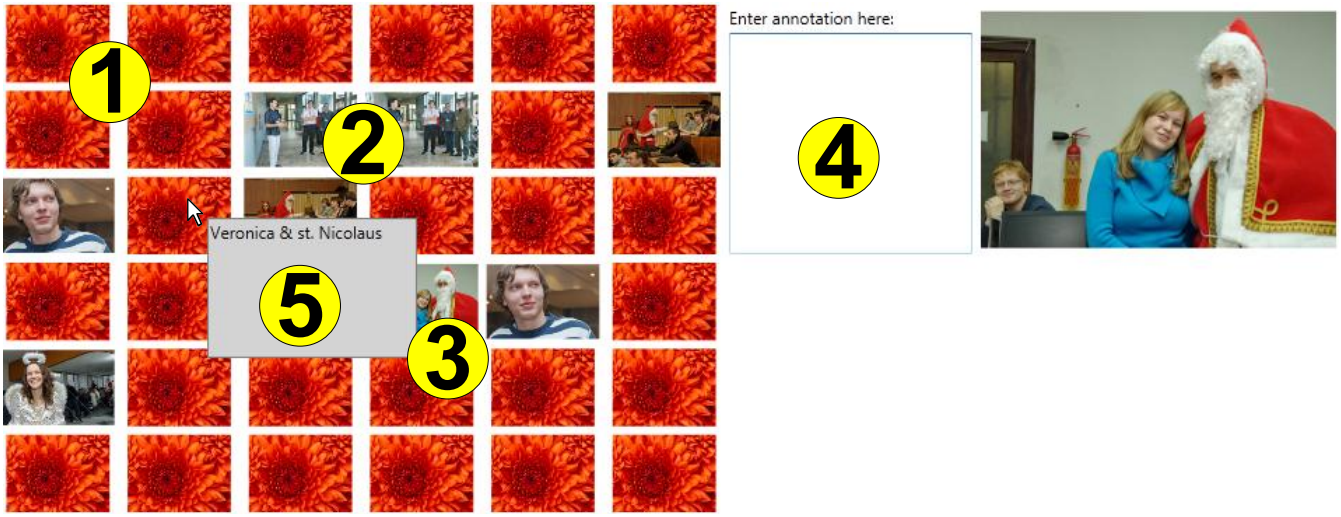

Figure 1: The PexAce game's player interface. Left grid represents game board with cards (1), some pairs already disclosed (2), one currently flipped (3), annotable via text field (4). Hovering over undisclosed cards yields annotation tooltips inserted by player in the course of game actions (5).

players annotating the same image. With the low number of players however (a case of using the game for personal image annotation), the method is prone to let only few tags trough filter, rendering the approach inefficient. Other heuristics were therefore devised.

- If the same user repeats the same suggestion multiple times (we used 2) on the same image. Sometimes, players encounter the same image in different game sessions and provide annotations containing common terms. Repetition might indicate (considering also the player's motivation) that the term is relevant to the image, even it was not confirmed by other player.

- If there is a suggestion used by the same player over multiple images (we used 4) in the same album. Here, the heuristics counts on a widespread practice of image owners to organize their images into album directories. It assumes that a concept or term might recur within multiple images in an album or even in the whole image set. In particular, this might be case for some owner-specific information like person names.

4. EXPERIMENTS

The purpose of our experiments and associated study was to explore the capabilities of our method for personal images metadata acquisition. We conducted it as a combination of qualitative and quantitative methods. Quantitatively, we measured correctness, specificity and understandability of tags acquired through our method. Qualitative study examined further characteristics of the gameplay and of produced metadata and was conducted as an interview, using a similar methodology as used by Vainio et al. [13].

Hypotheses and questions. Assuming a gameplay over personal images and low number of players

1. Do tags generally describe the images involved in the game?
2. Do tags address the specific context of the social group they belong to?
3. How is the performance of the tagging affected, if players are aware about the true purpose of the game (e.g. annotation of their own images)?
4. What types of group-specific metadata players express through their annotations (e.g. person names, places, events)?
5. Over which types of images, the approach works (e.g. portraits, groups, situational images, no-persons images)?

Participants. The study was conducted with 8 participants split into two groups. In each group, participants belonged to a same social circle, having common interests and also images they were familiar with. In each group, two participants were designated as players, one as a judge for tag evaluation and one helped with preparation of image data set.

Data. For each group, a set of 48 images was created. Images were drawn from three albums belonging to the group with the help of the group member. Each image was categorized either as *portrait*, *group*, *situational* (may involve persons, but has a strong dynamics in it) and *other/non-person* (e.g. architecture, landscapes, things). The albums and images were selected in a manner that from each album, an equal number of images was drawn to each category.

Methodology. The game experiment and study was done controlled in the form of interview. The interviews were done individually. We strictly followed prepared scenarios for each type of participant. The players were explained the rules and features of the game. Prior to that, in one group, they were also introduced (as in Vainio's study) to the concept of image metadata and also about the true purpose of the game – creating annotations for themselves. This was done, so we could measure the impact of this knowledge with reference to the other group. All the players were

free to speak any time during the gameplay and after it, not just to answer our questions. These expressions were evaluated later, qualitatively.

As for the judges, they were introduced to the concept of (personal) image metadata, but not about the game. We did this to keep judges as objective in the validation process as possible.

Process. The player interviews with gameplay were executed. As a language, we used Slovak. During them, each player played three games (with board sizes 6x6, 8x8 and 10x10), thus annotating each image twice (in 10x10, the two pairs remaining for 48 prepared pairs has been drawn randomly and were not considered in the experiment). After all games have been played, the tag extraction procedures were run. Then, the judges evaluated each tag assignment for their group, answering questions:

1. Is the tag *correctly describing* the image?
2. If it is a correct tag, is it also *specific* for the group (e.g. probably not discoverable by non-group player)?
3. If it is a correct tag, is it *understandable* for a non-group member (e.g. for a portrait image an assigned name is a self-explanatory and understandable tag)?
4. If this is a specific tag, of which *type* is it (choose one of the followings: a person name, a place name, an event where a picture was taken, other)?

4.1 Experiment results

Throughout the experiment, a total number of 366 tags was extracted using the three heuristics mentioned above. The "traditional" *cross-player-voting* heuristics (used in also in the original PexAce) yielded (as was expected) just one third of this number (exactly 122 tags). The rest was identified by the second (196) and third (48) heuristics (in that order, i.e. already identified tags were skipped by latter methods). This shows a *major increase of tag output quantity* for our extension of the original PexAce game – a traditional validation heuristic was weak, but the use of less restrictive heuristics paid off and increased the tag gain.

The quantitative results of our experiment are summarized in the Table 4.1. For both groups together, the correctness of tags is about 90% (96% resp. 84%) and therefore we consider our approach as able to acquire valid tags. In average, about 38,5% (44% resp. 33%) of correct tags were social-circle-specific, so our approach is also able to produce tags valuable for personal archives.

Table 1: Table showing summary results of image tag evaluation (correctness, specificness for the social group and understandability by group non-member) for image sets of both social groups (aware or unaware of the game's purpose).

	Aware (253 tags)			**Unaware (108 tags)**		
	Corr.	Spec.	Und.	Corr.	Spec.	Und.
Portraits	0,98	0,61	0,71	0,77	0,53	0,87
Groups	0,97	0,57	0,74	0,76	0,45	1,00
Situations	0,92	0,41	0,77	0,93	0,19	1,00
Other	0,98	0,18	0,82	0,88	0,15	1,00
Average	**0,96**	**0,44**	**0,76**	**0,84**	**0,33**	**0,97**

Considering the total number of tags produced by each group and the relative correctness, resp. specificity of tags, we can see that though both groups provided some value. The group where players were "initiated" to the purpose of the game produced significantly better tags in absolute quantity (total number of tags produced by the "initiated" group was 2.5 times larger) and relative quality. Only the understandability factor is reverse, which we explain to be a consequence of higher absolute number of specific tags passing through and having exclusive meaning (for instance, there was a tag carrying name of event related to images, not known to people who did not participated). More or less, we can conclude that awareness about the purpose significantly improves our method's results, however, as our qualitative study has shown, players were also less enthusiastic about gameplay when they were aware of the purpose.

According to judges, the specific tags mostly involved person names (53%). In lesser counts events (21%) or places (15%) were present too. The rest (11%) was categorized as "other" and involved mostly features originating from humor related annotations that players were using to entertain themselves. The humor is also a possible difference between performance of social groups in terms of specificness of tags for the *situational* type of images. In the "not initiated group", the humor was present within the player annotations and caused that term matching was not so successful for specific tags (with the far more disciplined "initiated group" on the other hand). Both groups (as was expected) performed well with *portrait* and *groups* types and were quite unsuccessful with *non-person (other)* type of images.

As a part of the qualitative study, we recorded that both groups enjoyed playing over their own images, particularly with those, they have not seen for longer time. A possible negative feature of the game was also detected: the players were confused and skeptical about having the same image in more than a one game shortly afterward (which happened because of the design of our experiment). They reported, that they tended to use exactly the same text for annotation in both cases (which could potentially harm the term extraction process) and that they were not sure, whether they had seen the image in the current or a previous game. A future-work image picking algorithm would have to take into account the last date-of-see for each image and player.

5. CONCLUSION

Based on the experiments we conducted, we consider our game-based approach as *capable of delivering metadata suitable for personal image archives.* The main advantage of our approach is, that it can work with limited number of participating players. This makes the game *suitable to deploy within small social-circles* or ultimately also *for individual users.* The small number of players is substituted by unique combination of player motivation to provide *valid annotations* – the game and competition experience as well as promise of working on one's own metadata (resp. image retrieval capabilities). We further shown that:

- The quantity of tags gained by our method was three times higher, than it would be if we only use the cross-player-voting (the original PexAce approach). Our method is therefore capable of delivering tags in higher quantities while retaining a comparable quality of tags (even if we assume, that the cross-player-voting con-

tributes the 100% quality, the latter extraction heuristics would have at least 85% output correctness over the rest of tags).

- The difference caused by increased motivation and awareness of the game's purpose have (in contrast with other GWAP-related works stating the exact opposite [10]) significant positive impact on the overall approach performance in terms of tag quality and quantity. A minor drawback however, was the lesser enthusiasm for playing in the (purpose) "initiated" group.

- Majority of specific tags produced carry the names of persons and lesser the names of places or events.

- The game delivers correct tags for all image categories, but in case of images without any persons, the number of social-circle-specific tags is low, which renders our approach limited for this type of images.

For the future work, we consider further enhancements to our method. From the implementation standpoint, we plan to construct and deploy a web-based version of the game, which would be able to work with player's online albums and repositories (e.g. Flickr, Zenfolio). Next, we want to move toward more advanced methods of term extraction: we plan to employ term synsets (for term cross-validation) and NER techniques, both to improve the valid term gain from the same number of annotations.

Acknowledgment

This work was partially supported by the grants VG1/0675/11/2011-2014, APVV-0208-10 and it is the partial result of the Research & Development Operational Programme for the project Research of methods for acquisition, analysis and personalized conveying of information and knowledge, ITMS 26240220039, co-funded by the ERDF.

6. REFERENCES

[1] M. Bieliková and E. Kuric. Automatic image annotation using global and local features. In *Proc. of the 2011 6th Int. Work. on Semantic Media Adaptation and Personalization*, SMAP '11, pages 33–38, Washington, DC, USA, 2011. IEEE CS.

[2] E. Chang, K. Goh, G. Sychay, and G. Wu. Cbsa: Content-based soft annotation for multimodal image retrieval using bayes point machines. *IEEE Transactions on Circuits and Systems for Video Technology*, 13:26–38, 2003.

[3] C. Cusano, G. Ciocca, and R. Schettini. Image annotation using svm. In *Proc. of internet imaging*, pages 330–338, Bellingham, 2004. V.SPIE.

[4] P. Duygulu, K. Barnard, J. F. G. d. Freitas, and D. A. Forsyth. Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary. In *Proc. of the 7th European Conf. on Computer Vision-Part IV*, ECCV '02, pages 97–112, London, UK, UK, 2002. Springer-Verlag.

[5] K. Eduard. Automatic photo annotation based on visual content analysis. *Information Sciences and Technologies Bulletin of the ACM Slovakia*, 3(2):72–75, 2011.

[6] F. Fauzi, J.-L. Hong, and M. Belkhatir. Webpage segmentation for extracting images and their surrounding contextual information. In *Proc. of the 17th ACM Int. Conf. on Multimedia*, MM '09, pages 649–652, New York, NY, USA, 2009. ACM.

[7] R. Gligorov, M. Hildebrand, J. van Ossenbruggen, G. Schreiber, and L. Aroyo. On the role of user-generated metadata in audio visual collections. In *Proc. of the sixth Int. Conf. on Knowledge capture*, K-CAP '11, pages 145–152, NY, USA, 2011. ACM.

[8] R. Guermazi, M. Hammami, and A. Ben Hamadou. Classification of violent web images using context based analysis. In *Proc. of the 2010 ACM Symposium on Applied Computing*, SAC '10, pages 1768–1773, New York, NY, USA, 2010. ACM.

[9] C.-J. Ho, T.-H. Chang, J.-C. Lee, J. Y.-j. Hsu, and K.-T. Chen. Kisskissban: a competitive human computation game for image annotation. In *Proc. of the ACM SIGKDD Work. on Human Computation*, pages 11–14, New York, NY, USA, 2009. ACM.

[10] M. Krause, A. Takhtamysheva, M. Wittstock, and R. Malaka. Frontiers of a paradigm: exploring human computation with digital games. In *Proc. of the ACM SIGKDD Work. on Human Computation*, pages 22–25, New York, NY, USA, 2010. ACM.

[11] V. Lavrenko, R. Manmatha, and J. Jeon. A model for learning the semantics of pictures. In *Advances in Neural Information Processing Systems 16 (NIPS)*, pages 553–560. MIT Press, 2003.

[12] B. Nagy. Acquisition of semantic metadata via an interactive game. In *Personalized Web - Science, Technologies and Engineering : 9th Spring 2011 PeWe Work. Proc.* Slovak University of Technology, 2011. [pewe.fiit.stuba.sk/ontoparty-2010-2011-spring/].

[13] T. Vainio, K. Väänänen-Vainio-Mattila, A. Kaakinen, T. Kärkkäinen, and J. Lehikoinen. User needs for metadata management in mobile multimedia content services. In *Proc. of the 6th Int. Conf. on Mobile Technology, Application & Systems*, Mobility '09, pages 51:1–51:8, New York, NY, USA, 2009. ACM.

[14] L. von Ahn and L. Dabbish. Designing games with a purpose. *Communications of the ACM*, 51(8), 2008.

[15] J. Šimko and M. Bieliková. Games with a purpose: User generated valid metadata for personal archives. In *Proc. of the 2011 Sixth Int. Work. on Semantic Media Adaptation and Personalization*, SMAP '11, pages 45–50, Washington, DC, USA, 2011. IEEE Computer Society.

[16] J. Šimko, M. Tvarožek, and M. Bieliková. Semantics discovery via human computation games. *Int. Journal on Sem. Web and Inf. Systems*, 7(3):23–45, 2011.

[17] L. Wu, L. Yang, N. Yu, and X.-S. Hua. Learning to tag. In *Proc. of the 18th Int. Conf. on World wide web*, WWW '09, pages 361–370, New York, NY, USA, 2009. ACM.

[18] V. Zsombori, M. Frantzis, R. L. Guimaraes, M. F. Ursu, P. Cesar, I. Kegel, R. Craigie, and D. C. Bulterman. Automatic generation of video narratives from shared ugc. In *Proc. of the 22nd ACM Conf. on Hypertext and hypermedia*, HT '11, pages 325–334, New York, NY, USA, 2011. ACM.