

High Quality Photo Collection via Gamification

Anca Livia Radu
DISI, University of Trento, Italy
ancalivia.radu@unitn.it
Bogdan Ionescu
LAPI, University Politehnica of
Bucharest, Romania
bionescu@imag.pub.ro

Aliaksandr Autayeu
DISI, University of Trento, Italy
autayeu@disi.unitn.it
Fausto Giunchiglia
DISI, University of Trento, Italy
fausto@disi.unitn.it

Abstract

This paper addresses the issue of gathering high quality photographic representations of the world by acquiring representative and diverse images with locations. These locations refer to different public spaces such as parks, museums, administrations, monuments, etc. The representations should be clear and focused so anyone seeing them will easily understand what they depict. This work comes in the context when the typical users are interested in receiving accurate relevant-to-the-query and non-redundant images so they can build a correct exhaustive perception of the query. We propose to tackle this issue by combining two approaches previously considered disjoint: gamification, i.e., involving people for photo generation and media analysis techniques for participants reputation assessment and photo filtering. This paper presents our preliminary results achieved for implementing the gamification strategy.

1 Introduction

The online image collections increased continuously and the multimedia resources are various. As a global effort, the web is already populated with many images of entities in general and locations in particular. However, there are not enough high quality images to represent the uniqueness of these entities [1].

Copyright © 2015 for the individual papers by the paper's authors. Copying permitted for private and academic purposes. This volume is published and copyrighted by its editors.

In: F. Hopfgartner, G. Kazai, U. Kruschwitz, and M. Meder (eds.): Proceedings of the GamifIR'15 Workshop, Vienna, Austria, 29-March-2015, published at <http://ceur-ws.org>

Further, the amount of low quality images on the social online platforms like *Panoramio*¹, *Picasa*² or *Flickr*³ is overwhelming so that even effective retrieval techniques have difficulties to filter them out. Furthermore, the current retrieval techniques are also limited since they rely mainly on text, image, or more recently on GPS coordinates. Textual tags tend to be noisy or inaccurate, automatic content descriptors fail to provide high-level understanding of the scene while GPS coordinates capture the position of the photographer and not necessarily the position of the query object. This limits significantly the quality of the retrieval results [2].

The current solutions combine image retrieval from the online platforms with media analysis techniques. These techniques are limited since semantic notions cannot be efficiently translated for media. Moreover, these solutions aiming at improving the retrieval focused until recently mainly on the relevance. However, an efficient system should provide results that are relevant, but that also cover different aspects of the query and not duplicates of the same perspective, so the user can build a correct and complete view of the query. In this paper we focus more on the diversity issue.

Research on media analysis reached the point where quality photo collection requires the use of user expertise. Thus, we try to overcome these shortcomings, by proposing a new hybrid mechanism combining two fields: gamification as human intelligence and proved-to-work media analysis techniques.

The remainder of the paper is organized as follows: Section 2 summarises the related work. Section 3 details our contribution as a hybrid mechanism which implies gamification for image generation and then automated media analysis techniques for computing the reputation of the

¹<http://panoramio.com>

²<http://picasaweb.google.com>

³<http://flickr.com>

players. Section 4.1 presents the studies conducted for evaluating the structure of the gamification part and also the definition of representative and diverse images set. In Section 4, the first pilot test of the photo hunting festival is presented and the results we obtained. Section 5 concludes the paper.

2 State of the art

Various approaches focused on improving search capabilities on the current social media platforms and on providing users both representative and diverse results. Some of the most successful approaches on this direction are related to re-ranking [3].

Re-ranking techniques attempt to refine the initial results retrieved from the online platforms using the visual content of the images. An example is the approach in [4]. The paper defines a retrieved image as being representative and diverse if it is representative for a local group in the set, it covers many distinct groups and incorporates an arbitrary pre-specified ranking as prior knowledge. To determine these properties, authors propose a unified framework of absorbing Markov chain random walks. A different approach [5] estimates the relevance scores of images with respect to the query based on both the visual information of images and the semantic information of associated tags. Further, it estimates the semantic similarities of the images based on their tags. Using the two estimations, the ranking list is generated by a greedy ordering algorithm which optimizes average diverse precision. However, these techniques are limited, because they are not capable of “understanding” the world and the content of the images, since automatic content descriptors fail to provide high-level understanding of images.

Gamification is another field with potential, but without concrete approaches in collecting representative and diverse images. Games with a purpose are a narrow example of gamified activities where humans become part of collective computation. They are used in different areas, including computer vision and Internet retrieval. For example, *ESP Game* aims at collecting correct labels for images. Two players are asked to provide the same word or phrase without communicating while an image is the only thing they have in common. The only modality of providing the same output is to write down a correct label to the image [6].

In time, games evolved and now different location-based games with a purpose are also used for content generation. *Urbanopoly* [7] is a social mobile and location-based game with a purpose inspired from the famous Monopoly real board game. It aims at collecting, verifying and correcting data about “venues” in urban environment. As in the original version, the players are landlords that aim to collect as many venues as possible. But in the case of *Urbanopoly*, the venues are real places in the surroundings of the play-

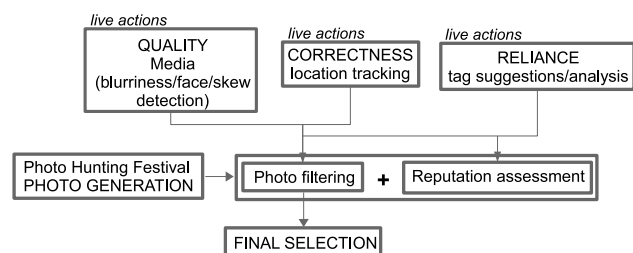


Figure 1: Proposed hybrid mechanism.

ers, taken from *Open Street Map*⁴.

Each user has an initial budget to buy venues. An available venue can be bought by a user with enough money by simply paying and providing some information about the venue (name and category). If the venue is not available, the user can spin a “wheel of fortune”. The possible outcomes are: money, request to provide a poster for that venue (picture and information) in exchange for money or request to verify your own venue’s poster provided by another player.

Some location-based applications like *Foursquare*⁵ and *Brightkite* have included the process of collecting urban information and images with real world locations [8]. Users can check-in to the service with a status and a location and establish friend relationships to view each other’s updates. These applications help people find out when friends are nearby or to decide where to go out.

A photographic marathon is a fun and challenging competition organised for collecting images with locations and events. For example, *Maratona Fotografica*⁶ attracts people for collecting images of historical centres of different Italian cities. Participants receive a list with locations to reach and they are rewarded for the fastest tour or the most interesting set of pictures.

Photo contest sites became more and more popular due to fast development of technology, when almost all of us have a camera in our pocket. *ViewBug*⁷, *Pixoto*⁸ or *PHOTOBraniac*⁹ are three examples of photo contest sites where participants can win cash and organizers collect images of different types.

Re-ranking techniques are incapable of understanding the content of the images and thus are limited in refining the results retrieved from online platforms. In contrast, our approach takes advantage of the collective human intelligence through gamification for quality photo generation and further applies only proven-to-work automated techniques.

⁴<http://www.openstreetmap.org/>

⁵<http://www.foursquare.com>

⁶<http://www.disturbo.net/>

⁷<http://www.viewbug.com/>

⁸<http://www.pixoto.com//>

⁹<http://www.photobrainiac.com/>

3 A photo hunting festival

In this paper we address the problem of constructing a large database with high quality images of locations, both representative and diverse. Our approach is designed to avoid the limitations of the online platforms and re-ranking techniques by combining two fields previously considered disjoint:

1. gamification: human expertise employed for improved photo generation;

2. automated media analysis with location tracking and tags analysis used for filtering the images and assessing the reputation of the participants. More precisely, automated media analysis plays the role of detecting undesired images: low quality (e.g., blurred or unfocused), containing faces in focus and duplicate images.

2. automated media analysis with location tracking and tags analysis that support gamification in returning quality results by filtering the images and assessing the reputation of the participants. More precisely, automated media analysis plays the role of detecting undesired images: low quality (e.g., blurred or unfocused), containing faces in focus and duplicate images.

The gamification part of the hybrid mechanism was implemented and thus tested in a preliminary validation iteration. The second part is to be implemented in the next larger iterations.

3.1 Structure of the festival

The paper uses the following concepts and notations:

– *Festival*: event that bundles together different sub-events sharing the same organization and the same period and area.

– *Challenge*: sub-division of the festival that stands as a separate competition and consists of an individualized list of items that participants need to solve.

– *Item*: description of a location that participants are required to capture through more pictures.

– *Checked item*: an item for which the participants completed all the requirements (i.e., the route of the participants includes the location of that item and they also provided pictures of it).

– *Representative and diverse images*: A photo is considered to be representative for the location if it depicts it in a clear, distinctive and focused manner the uniqueness of that location. Anybody knowing the appearance of the location and seeing the image should easily recognize it.

Other images are also accepted as being representative: images taken in different moments of the year (long-term temporal variation) and/or under different weather conditions; images taken during different moments of the day (short-term temporal variation); edited images (creative

views) as the meaning of the image is not affected by the manipulation and its only purpose is to improve the appearance of the image.

A set of images is considered to be diverse if each image depicts different visual characteristics of the target with a certain degree of complementarity. The whole obtained from the entire set of images creates a complete, clear and meaningful vision of that specific location.

The diversity focuses only on the aspects that directly refer to the location itself. Therefore, even if light manipulations and temporal variations (dawn-day-dusk-night and spring-summer-autumn-winter) with their corresponding weather conditions are accepted and contribute to building a large overview of the location's background, they are not sufficient for building a complete overview over the concrete location itself. Instead, in order to build a complete vision of the location, the following variations are accepted: outside/inside, frontal/side/back, top/bottom, large/narrow, distant/ close, entire/partial views and also typical aspects that are specific to that certain location, e.g., objects, paintings, animals, plants, etc. from the inside or outside of the location.

– i_k : one given item k in the challenge chosen by the team;

– q^{i_k} : the number of quality images uploaded by the team for a certain item i_k ;

– p^{i_k} : the number of images uploaded by the team for item i_k ;

– I : the number of items in the challenge chosen by the team;

– r^{i_k} : the number of reliable images uploaded by the team for a certain item i_k ;

– c^{i_k} : the number of correct images uploaded by the team for a certain item i_k .

We propose the implementation of the gamification part as a photo hunting festival where participants have the possibility to choose among several challenges. We are interested in attracting different types of audience. Therefore, each challenge covers various activities and requires different skills. The simplest challenges require participants to go to certain locations in the city and take pictures. The most difficult ones add a degree of creativity, knowledge of history and physical skills.

Motivating people to participate at the festival is done both implicitly through the photo taking action, the mechanics of the festival and the social aspect (getting together, competing, entertainment, etc.), but also explicitly using clear incentives. The latter is achieved through a master class on photography that is held before the opening of the festival by an expert team. They will explain the participants how to take quality images with different devices and also which are the expectations regarding their images.

All challenges run against time (typically, 2-3 hours). Also, the teams have the freedom of creating their own path, as long as they approach and capture the required locations. The physical approach of locations is verified u-

sing a smartphone application that participants are required to install prior to the beginning of the challenge. All challenges are considered complete only if a given number of locations in the list are captured (e.g., 9 out of 10 locations).

The challenges running in the festival are:

1. Walking Challenge – can be seen as urban hiking. Participants are given a list with the exact name and address of different locations to take pictures of. No vehicle is allowed.

2. Creativity Challenge – participants will receive a list of riddles with locations. Neither the name, nor the address of the locations will be provided. They are further asked to guess the locations behind the riddles and show their answers through physical approach and pictures of those locations. No vehicle is allowed.

3. Complete the Picture Challenge – participants will receive a list of photo-riddles. A photo-riddle is an incomplete picture from which the participants must figure out the concrete location and take other pictures. For example, they have to take the whole picture given a certain part, to take the right part given the left, take the front given the back, etc. Neither the name, nor the address of the locations will be provided. No vehicle is allowed.

While Creativity and Complete the Picture Challenges are active, the participants have also the possibility to create and upload riddles for the other participants. In order to ensure that the riddles uploaded by the participants are correct and solvable, the organizers check them before releasing. However, in case the number of riddles uploaded by the participants cannot be handled by the organizers, they will be uploaded unchecked to the application. Participants will be warned about the unchecked riddles.

4. Time Travel Adventure Challenge – participants will be provided historical pictures with locations as photo-clues and a time-slot referring to the building time. Teams will have to recognize the locations and to bring new ones. Neither the name, nor the address of the locations are provided. No vehicle is allowed.

5. Bike Challenge – is similar to the walking challenge, but the participants are allowed to use bicycles, as a non-polluting vehicle. The locations cover larger areas compared to the other challenges.

6. Journalism Challenge – allows participants to act like observers and take pictures of the most important moments and aspects of the festival. Part of these pictures will be used for advertising the event. Funny pictures, with as many participants as possible are favoured.

3.2 Players' reputation

A large number of participants and the resulting competition may attract low quality results and inadequate behaviour, like cheating. Given both our purpose of gathering high quality diverse photos and the expected large amount of user generated content, it is mandatory to find a mechanism to moderate the festival. The moderation refers to assessing the quality of the uploads and further pre-filtering them and computing the reputation of the participants. In our case, the entire process is performed using the automated media analysis as described in Section 3.

Thus, the criteria that we are searching for in the teams' uploads that further allow us to extend the reputation of the participants are *quality*, *correctness* and *reliance*. Apart from contributing to the reputation computation, those images that do not follow the criteria of quality and correctness are automatically filtered out.

Quality in images, as computed in equation (1), refers to a set of requirements with a big impact on images appearance. The requirements relate to the quality of the image itself, disregarding the content significance: sharpness, alignment, no faces in focus. Thus, we accept only sharp, aligned and with-no-faces-in-focus images.

$$quality = \frac{1}{I} \sum_i \frac{qi_{i_k}}{pi_{i_k}} \quad (1)$$

The value of *correctness* as computed in equation 2 judges the capture of the images during the challenges. The participants might try to cheat by uploading images that do not belong to them. Thus, we verify both the list of locations approached by the teams and the EXIF in the photos. In case they upload an image with a certain location they “pretend” to have reached, but which does not belong to their tracked path, it means they are trying to “game” the challenge.

$$correctness = \frac{1}{I} \sum_i \frac{ci_{i_k}}{pi_{i_k}} \quad (2)$$

Asking participants to tag the images they provide is a modality of collecting additional information and, also, judging the *reliance* of their images (see equation 3). Thus, for each location, the teams are asked to judge a series of mandatory suggested tags and further to freely add, if they want, some extra-tags. The mandatory tags describe the name of a location and will be provided according to the position of a team (recorded by the smartphone application). They are of two types:

– checked tags: tags specifically chosen incorrect (i.e., describing a location that is obviously too far from the actual position of the team). They are meant to test the credibility of the team and will not be attached to the images.

– unchecked tags: tags describing a location close to the actual position of the team. However, the localization system is sometimes not accurate enough to distinguish between

locations close one to each other. Thus, the team might need to select one tag among several located in the same small area. The tags provided by the teams will or will not be accepted, according to how well they succeed to filter the checked tags.

$$reliance = \frac{1}{I} \sum_i \frac{r_{i_k}}{p_{i_k}} \quad (3)$$

All three variables range from 0 to 1. The three variables have equal contribution in computing the reputation of each team:

$$reputation = (quality + correctness + reliance) / 3 \quad (4)$$

4 Validation results

4.1 Validation of festival's structure

Optimal results collected through the festival require a fine structure tuning. Thus, we conducted a study that helps to identify and build the best structure for each challenge and of the entire festival.

Another study was conducted with the purpose of evaluating and improving the representativeness and diversity definitions presented in Section 3.1.

Both studies were conducted between 24th September and 24th October 2014 and they were distributed online to 34 people. In total, 26 contributors responded for the first study and 25 for the second one. All of them were located in Romania.

The questions requiring the contributors' agreement toward a certain idea/definition use a 7 points Likert scale with the options: "Strongly disagree", "Disagree", "Mildly disagree", "Neither agree, nor disagree", "Mildly agree", "Agree", "Strongly agree".

For the first study, contributors were initially asked to present their level of familiarity with the concept of scavenger hunt. Results indicate that the majority (42.30%) of contributors never heard of scavenger hunt. Further, contributors either have heard, but know nothing about it (15.38%), or know something, but never participated (26.92%). Finally, only 15.38% of the contributors participated to a scavenger hunt.

Further, we asked the contributors about their level of agreement towards different scenario elements, like the idea of team participation (50% Agree), photography tutorial before the opening of the festival (69.23% Strongly Agree) and also for the installation of a smartphone application prior the beginning of the festival (50% Agree that the application is useful).

Further, we presented each of the six available challenges for which we asked them to rate to which extent they would be interested to participate. Thus, they showed their interest in competing first at Time Travel Adventure



1) 1893 2) 1923 3) 1864

Figure 2: Items for Time Travel Adventure Challenge.

Challenge (46.15%), then Walking, Creativity, Bike Challenges (30.76%), further to Complete the Picture Challenge (26.92%) and lastly to Journalism Challenge (3.84%). In the same time, we collected the contributions of the participants that had the possibility to provide ideas about new challenges or to modify the current ones.

Finally, contributors were questioned about the additional activity of creating new riddles for other participants in Creativity and Complete the Picture Challenges. Thus, 46.15% of the contributors for the first, respectively 57.69% for the second one declared this activity makes them wanting more to participate.

For the second study, a majority of 56% agreed with the definition of representativeness, while 56% strongly agree with the diversity.

4.2 Description of the implementation

A first pilot test was organized in Bucharest, Romania for testing all the requirements and problems that may occur in large scale festivals. In this iteration we only implemented the gamification part from the hybrid approach described in Section 3. Thus, we launched 3 out of 6 challenges as presented in Section 3.1: Walking, Creativity and Time Travel Adventure Challenges (see samples below).

Each of the three challenges contained 7 items, thus a total of 21 locations to be captured. The locations referred to different museums, administrative buildings, monuments, squares, etc. from Bucharest, more or less known to the public. 7 mixed teams (individual or with 2 people) participated at all three challenges after being presented a set of participation rules. Samples from the three challenges are presented below.

Challenge 1: Walking Challenge

1. Old Princely Court and Church (Palatul și Biserica Curtea Veche), str. Franceză 25-31.
2. The Linden Tree Inn (Hanul cu Tei), str. Lipscani 63-65.
3. Curtea Berarilor, str. Șelari 9-11, București.

Challenge 2: Creativity Challenge

1. Find a very touristic pedestrian street named after Leipzig.
2. Find the oldest "recently in use" hotel building in Bucharest that kept the initial 19th century inns structure, despite the repeated restorations.
3. Find "the City of the Cross", patronised by the saints that orthodoxism celebrates in one week from today.



Figure 3: a), b) *The Comedy Theatre*; c), d) *Curtea Berarilor*.

Table 1: Results collected per challenge.

	av. nb. of collected images	av. nb. of correct images
Challenge 1	8.3	8.2
Challenge 2	7.65	6.5
Challenge 3	7.91	7.53

Challenge 3: Time Travel Adventure Challenge

Figure 2 depicts items for the Time Travel Adventure Challenge (including the building year).

4.3 Results of the pilot test

During the festival, we collected a total number of 1170 images for all 21 locations. We want to check the motivation in the gamification part, before spending time on implementing the entire application and the media analysis techniques. Thus, for validation and analysis purposes we manually checked the results. We discovered a majority of 1089 images were correct. Apart from being both representative and diverse images, the images are of high quality, despite the non-professional status of the participants. The rest of 81 images were classified as being incorrect, not because they were not representative and diverse for the captured locations, but because it happened that the participants mixed some locations and they captured different locations than those required in the challenges. That happened partly because of their low experience with the locations and partly because of the difficulty and even ambiguity mostly in the Creativity and Time Travel Adventure Challenges.

As mentioned in Section 4.2, we asked the participants to collect a minimum number of 5 pictures per location. The results show an average number of 7.95 images per location, varying from 0 (for locations not reached by the teams) to 26 images.

Further, eliminating the locations that teams did not reach, the average number of collected images rises to 8.60. From this fact, we can conclude that participants were willing to capture a higher diversity than the required minimum. From all images, an average number of 7.41 correct images were collected per location. The rest of incorrect images comes from those locations confounded or not reached. Thus, we were successful in collecting the desired high quality images using our approach.

Further, Table 1 presents individually for each challenge both the average number of totally collected images and the

Table 2: Results per challenge.

	Walking	Creativity	Time Tr. Adv.
nb. of imgs	407	375	388
imgs/team/item	8.3	7.65	7.91
correct imgs/team/item	8.2	6.5	7.53
imgs/team/reached item	8.3	8.62	8.88
correct items	6.85	5.14	5.71

average number of correctly collected images.

As expected, challenges requiring additional skills (Creativity and Time Travel Adventure Challenges) collected a lower average number of images compared to the simpler ones (Walking Challenge). Moreover, the average number of correctly identified and captured location is 5.9 out of 7, given the fact that each of the three challenges was considered to be finished if 6 out of 7 locations were correctly captured. Additionally, Table 2 presents the results obtained individually by each of the three challenges.

To have a subjective measure of performance, two pairs of image-sets are illustrated in Figure 3 as a visual comparison between the results obtained through our approach (Figures 3a and 3c) and the first results retrieved from *Flickr*¹⁰ and *Google Images*¹¹ (Figures 3b and 3d) in January 18th 2015 using as keywords the names of two locations (*The Comedy Theatre Bucharest* and *Curtea Berarilor*) captured during the festival. The images collected through our Festival are of high quality, from both representative and diverse point of view. On the other hand, the retrieved images depicted in Figure 3b are diverse enough, but mostly obviously do not depict the query location. Instead, the images from Figure 3d are representative for the query location, but lack diversity.

5 Conclusions

In this paper we proposed and evaluated a first part of a replicable novel hybrid mechanism that combines a photo hunting festival for photo generation and media analysis techniques for photo filtering and assessing participants reputation. Different variations of the approach can be run on large scale even with online coordination. In this way, many images of high quality can be collected, without relying on the limited image retrieval and media a-

¹⁰<http://flickr.com>

¹¹<http://images.google.com/>

analysis techniques. The target are locations from all over the world, from very popular to less known for which there are not enough online resources. Close future work consists in organizing a large scale version of the festival in Trento (spring) accompanied by the automated media analysis technique.

Location Sharing Application, SIGCHI Conf. on Human Factors in Computing Systems, 2011.

6 Acknowledgments

This research was supported by the Romanian Sectoral Operational Programme POSDRU/89/1.5/S/62557 and the Collaborative Project Smart Society: “Hybrid and Diversity Aware Collective Adaptative Systems: When people Meet Machines to Build a Smarter Society” (<http://www.smart-society-project.eu/>), funded by the European Commission’s 7th Framework ICT Programme for Research and Technological Development under the Grant Agreement no. 600854.

References

- [1] F. Schroff, A. Criminisi, A. Zissermann. *Harvesting Image Databases from the Web*, IEEE Transactions on PAMI, 2011.
- [2] A. Dia Miron, B. Moisuc, J. Gensel, M. Villanova-Oliver, H. Martin. *Towards the semantic and context-aware management of mobile multimedia*, Journal MTAP 53(2):391-429, 2011.
- [3] B. Ionescu, A. Popescu, M. Lupu, A. L. Gîncă, H. Müller. *Retrieving Diverse Social Images at MediaEval 2014: Challenge, Dataset and Evaluation*, MediaEval 2014.
- [4] X. Zhu, A. Goldberg, J.V. Gael, D. Andrzejewski. *Improving Diversity in Ranking using Absorbing Random Walks*, Int. Conf. HLT-NAACL, 2007.
- [5] W. Meng, Y. Kuiyuan, H. Xian-Sheng, Z. Hong-Jiang. *Towards a Relevant and Diverse Search of Social Images*, IEEE Transactions on Multimedia 12(8):829-842, 2010.
- [6] L. von-Ahn. *Games with a purpose*, IEEE Computer, vol. 39, no. 6, pp. 92-94, 2006.
- [7] I. Celino, D. Cerizza, S. Contessa, D. Dell’Aglia, E. Valle, S. Fumeo. *Urbanopoly – a social and location-based game with a purpose to crowd-source your urban data*, ASE/IEEE Int. Conf. on Social Computing, USA, 2012.
- [8] J. Lindqvist, J. Cranshaw, J. Wiese, J. Zimmerman. *I’m the Mayor of My House: Examining Why People Use Foursquare - a Social-driven*