Serious Games for Large-Scale Image Sensing

Bruno Simões[1]([✉]), Michele Bianchi[2], Alberto Debiasi[1],
and Raffaele De Amicis[1]

[1] Graphitech, Trento, Italy
{bruno.simoes,alberto.debiasi,raffaele.de.amicis}@graphitech.it
[2] University of Trento, Trento, Italy

Abstract. The access to large-scale imagery datasets has been a signifi-
cant obstacle to the success of many applications in application domains
that range from 3D modelling to augmented reality, and from infrastruc-
ture inspection to urban planning. Although large collections of images
already exist, from sources such as Bing Maps, Google Street View, and
many photo-sharing sites, they are incomplete, inaccurate and expensive.
A solution to this problem could be to leverage on large end-user com-
munities to collaboratively acquire and share information about their
surroundings. In this paper, we outline some basic mechanics in serious
games that can be explored for the purpose of data collection. Addi-
tionally, we describe new ways of guiding players' actions towards the
purpose of our game – image and video crowdsourcing.

Keywords: Serious games methodologies · Participatory sensing

1 Introduction

The ubiquity of full-fledged sensing, computing, and communication devices like
smartphones is paving the wave for new perspectives on how to accomplish large-
scale sensing. The phenomenon of large-scale sensing, which is better known in
the literature as participatory sensing, is an approach to data collection and
analysis in which individuals and communities use their personal devices to
acquire and explore specific aspects of their surroundings. The number of appli-
cations is quite vast and can range from environmental issues to culture. Imagine,
for example, that we are capable of collecting geo-referred images and videos that
depict every single corner of a town, at different times of the day and year. Then
we could use our framework [1] to reconstruct any object (e.g. animals, buildings,
etc.) and even entire cities in 4D (a 3D model with a time dimension), with the
use of computer vision techniques. The potential of these 4D representations is
huge. We could create virtual tours that would enable users to visit or explore
new parts of a town without going there; in addition to experience it from any
location and angle, at any specific time of the day or the year (e.g. day, night,
winter, summer, etc.). It would enable us to preserve and experience moments
like when a building is torn down, repaired or constructed, outdoor festivals,
etc. This data could also be ingested by domain-specific applications in order to

support their users monitoring and analysing, for example, the infrastructure of a building, for urban planning, or simply for the benefit of individual creators or creative industries (e.g. operating in 3D game content creation).

There are a couple of online services that provide access to large structured (e.g., Bing Maps, Google Street View, etc) and unstructured collections of images (Flickr, Picasa, Panoramio, Facebook). However, these repositories have several drawbacks if used in the aforementioned context. One the one hand, acquiring imagery from structured repositories can be financially expensive – structured repositories are often the core business of commercial enterprises. Additionally, data is only available in certain areas or it captures only certain features of the environment, e.g. streets. Another issue is that these repositories are updated very sporadically, thus, these collections may be out of date at some point; similarly, these collections maintain only a specific moment of day or of the year – e.g. when the last satellite picture was taken. On the other hand, we can find many unstructured repositories that are maintained by thousands of people that walk around everyday with devices that can take pictures or record videos. The only flaw in these services is that only the most attractive facades of landmarks are well-represented, against everything else that is very sparsely captured.

In our applicational context, this limitation represents a serious issue because to create 4D models of complex scenes or objects (e.g. a concert or the building appearance spanning across many years) we need to find an sustainable approach that enable us to collect the vast number of 'missing' view angles, along the time span we want to reconstruct, and without any special equipment. So far, many crowdsourcing applications have been developed. However, crowdsourcing applications have first to overcome the challenge of how to motivate massive crowds to participate. There are two main strategies that are widely used to influence user behaviour towards doing a specific action: serious games and gamification. Serious games are "games that do not have entertainment, enjoyment or fun as their primary purpose" [2,3]. Gamification is primarily characterised by the fact that the game, if defined, is always secondary to the tasks that have to be performed, e.g. adding a points system, peer pressure, leaderboards, as well as other things that normally would not be considered. Effective gamification exploits the user context to provide motivation specific to the situation, instead of simple integration of badges and leaderboards [4,5].

In this paper, we introduce a serious game to crowdsource the gathering of videos and pictures. Our key contribution is the game design that is proposed. In our game, the simple action of taking pictures serves two purposes: players perceive it as a mechanism to capture magical creatures; from our perspective it is a strategy to gather pictures. The outcome of the game, which is transparent to the gameplay, is a set of images spanning over large areas, from multiple angles, and a set of 4D models reconstructed from those images.

2   Related Work

The idea behind Serious Games (SG) is to design them so they provide entertainment to the players while primarily serving a greater purpose. The purpose of

the game can be to generate useful data, to communicate information, to teach concept. Most serious games are based on the work of von Ahn, who developed together with other authors the ESP game [6] and Peekaboom [7] to demonstrate the concept, along with other widely used games [8,9]. The ESP game was designed for the purpose of labelling random Internet pictures. The game is grounded on the idea that whenever two players visualise the same image they should come up with textual tags that match. In this game, players are paired randomly without any means of communication; therefore to earn points, they must find tags that most people would associate to the image. Peekaboom, on the other hand, involves defining the location of objects.

Foldit [10] is an experiment that serves to exemplify the relevance and efficiency of using games that take advantage of human' innate spatial reasoning abilities, to solve problems that computers fail to resolve. In this game, people can help research scientists to solve a protein-folding problem that had baffled them for more than a decade. The objective of the game is to find, for each digital 3D protein structure, the most tightly packed configuration.

Astro Drone [11] is another crowdsourcing effort based on data collection through a game. The aim of the game is to play a spaceship simulation game in which players have to control a spacecraft that has to fly close to a comet and then release a lander on it. The game sends the data extracted from the camera of the drone to ESA to be then used in the study of new automatic control systems (e.g. obstacles avoidance and docking).

PhotoCity [12] is a game with the purpose of collecting photos. This game relies on small groups of expert photographers highly motivated to acquire pictures that are useful to the reconstruction of buildings. Players are rewarded with castles and flags as they contribute to 3D replicas. The authors used vision techniques to reward only those players with useful input. The drawback of this strategy is that this game is designed to reward only those players with exceptional skills in photography, as a way to compensate for a limitation in their 3D reconstruction framework. Consequently, the game fails to implement a mechanism that can attract large communities and that can maintain them engaged independently of their skills or interests.

Finally, EyeSpy is another relevant game that proposed by Bell et al. [13]. The objective of this game is to collect pictures and tags that can be functional in a navigational context, for example, to give directions to someone based on landmarks that can be easily recognisable. To play the game, we just need to walk around, take pictures, and insert tags. Additionally, the player can geolocate and tag pictures collected by other players.

3   Methodology

Our game is split into three game moments: capturing creatures, exchanging cards and fighting other players. Capturing creatures requires players to go outdoors and use their smartphones in AR mode, see Sect. 3.4. There are two strategies for capturing creatures: take a picture to a spot that was identified directly

through the display or by "blindly" taking pictures around – to discover hidden gems. The game provides also a tool to assess the "elemental profile" of a physical location, which determines the probabilities of capturing a creature of a specific element and a radar for navigational purposes.

In this game, creatures are quite abundant in quantity but not in quality. As we will explain later, during game battles weak creatures will be absorbed by other creatures either to attack or to evolve. This mechanic was integrated to push players into taking pictures even if they do not need more creatures. It also generates dynamics that appeal to explorers user-type whom are looking to complete their list of creatures, and to killer user-type that want to gain the upper hand against others, e.g. influence points.

Fighting pits two players one against another for the supremacy of a guild over another, that is, for influence points. Players will have to use creatures they have captured in order to win. There are two different ways of engaging in a battle: (1) the player requests a online fight or (2) the game automatically initiates a fight between two players. The latter, called proximity battle, forces players to fight each other if they are capturing creatures in the same zone. Players can withdraw by paying a certain number of elemental cards or alternatively by sacrificing a creature.

A creature can be summon only after the DNA (picture) of the creature is uploaded to our backend infrastructure – an action that corresponds to reconstruct the creature from its DNA. We implemented this strategy of unlocking creatures at home or via WIFI to avoid exceeding their data plans. At the backend, we assess the veracity as well as the added value provided by the picture, which is a parameter in the computation of the set of abilities that is assigned to the creature.

In the next sections, we describe the most crucial gameplay elements in conjunction with the effect they should generate on the collection of pictures during the intervention.

3.1 Creatures' Distribution and Elemental Affinity

Creatures are aligned to one or more fictional elements that have been selected for the game. Each player character (selected during the registration process) has a stronger affinity to a pair of elements and a weaker affinity to a third element. Affinities change the probability of finding creatures of those elements, making it easier for elements with a stronger affinity and harder for elements with a weaker affinity. Nevertheless, this property does not change the fact that some elements are incompatible with certain environments. For example, fire-based creatures are unlikely to be spotted inside lakes, see Sect. 3.2 for more details.

To further push the interaction between players and to balance the way players can earn influence points, cities (physical space) and towns are automatically sub-divided into zones. The size of each zone in the spatial grid is inversely proportional to the number of pictures taken in the surrounding area, following the concept of a quad-tree. The purpose of this mechanism is to balance the spatial distribution of pictures and consequently to control the way players

can earn points. However, players shall perceive it differently: when a player captures a creature the control of that guild in that specific zone is increased (influence points). This new probability of finding a creature is presented as a decrease in the population of creatures (localised). The population is automatically restored over time, based on players behaviour or on external factors (e.g. an happening, a festival, etc.). Restoration over time has designed for continuous player involvement: newbies can always find a place to go for hunting, while veteran players are required to return regularly to their zones in order to maintain their local influence points. This strategy attracts not only the attention of killers and achievers player types [14,15], but also of achievers and explorers that are interested in finding rare gems where elements have not been depleted. The location and bearing of a picture within a zone are two of most relevant variables used to compute the probability of finding strong creatures, see Section. An internal algorithm was implemented to balance the scores by comparing these properties at local and global level – players always compete with someone.

3.2   Creatures' Habitat and Elemental Affinity

To maximise the realism of our game and to attract explorer-type players, we have implemented a service that constrains the probability of finding specific creatures in certain areas. Hence, marine creatures are most likely to be found nearby the sea and an elf nearby a forest. The game procedure is fully automatic, however, at the moment it supports only the European territory.
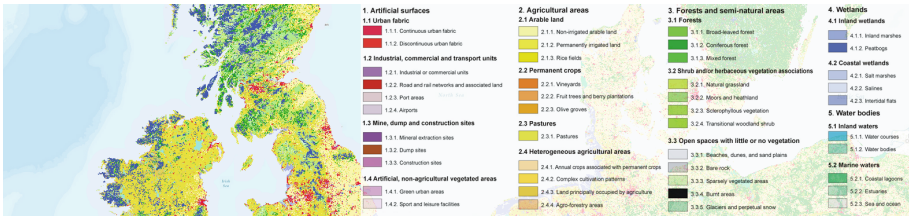

Fig. 1. Corine land cover.

The habitat mapping service is engineered as a two layers data service. The first data layer exposes the Corine Land Cover (CLC) of Europe as a service. The CLC nomenclature aggregates 44 land cover classes in a three-level hierarchy. Five main categories are "artificial surfaces", "agricultural areas", "forest and semi-natural areas", "wetlands" and "water bodies", see Fig. 1. The second data layer requests, to the first layer, the CLC code of a given location and then computes the list of compatible elements to that nomenclature.

The number of new creatures that is generated depends on the relevance of the area. Our system is capable of generating thousands of different creatures and descriptions without human intervention. During the creation process the system

takes into account the strategies described above. Every time a new card is generated, we have to update the database of the living creatures. See Sect. 3.4 for a few examples of automatically generated cards. Limited and thematic editions will also be considered.

3.3   Creatures Lifecycle

Given that many of our mechanics are based on probabilistic algorithms, both for generating and combining creatures, we have implemented an ageing effect – creatures will age and die – to prevent the use of singularities in fights. Creatures will age one year (creature's calendar) for each week (player's calendar). After a certain number of years ($\exists n \in \mathbb{N}, 2^n = i$, where i is the age of the card in weeks) the creature goes through an elemental evolution. At each evolution, the player can decide which element of the creature to level-up. A creature can live up to 110 years in average. Players can capture creatures with ages raging from 10 to 40 years old. Under special conditions, e.g. legendary creatures that correspond to high-value pictures, creatures can live up to 390 years.

The maximum lifespan of a creature is also shortened after each battle by one year. However, as a recompense, creatures evolve whenever the condition $\gcd(n, F_{n+1}) > 2$ is satisfied. The sum of all battles is represented by n; F is the Fibonacci Sequence and gcd is the greatest common divisor.

Fig. 2. Left: augmented reality interface. Right: collectible creatures.

Figure 2 depicts the interface to find creatures and market places, where creatures can be exchanged. By taking a picture the user accesses the closest and most central icon on the display.

3.4   Combat Mechanics

The game battle between players is split into 4 game moments: (1) summoning, (2) energy infusion, (3) energy release, and finally, (4) defence release. At the beginning of each turn – step 1 – players can summon new creatures into the battlefield. Players are free to pick the creature they want from their deck. The cost of the second phase – energy infusion – depends on the number and type of cards that are

on the battlefield: creatures in the battlefield will absorb cards from the player' deck until they collect the number (and type) of elements that are indicated in their properties. Creatures in the battlefield can also be sacrificed for this purpose. In a third phase – energy release – we take all creature' elements that are levelled-up as the player' magical attack. In step 4, the defender will have to infuse any missing elements (quality and quantity) from their deck to match the opponents attack. The first player to run out of cards looses the match as well as a percentage of the cards absorbed. The battlefield itself will have affinity to a few elements – that depends on the players' location or that is otherwise random – that changes the properties of an element to: double effect, no effect, or half effect.

Lastly, the game page features a leader board where players are ranked by influence points. There are five titles that players can earn: The Warlord (won the largest number of battles), The Expeditioner (possesses the largest number of gems), The Imperator (rules the largest number of areas), and The Evangelist (manage to convince the highest number of users to play the game). Local-level titles were also defined.

4   Conclusion

This paper describes the design of a serious game that aims at motivate players to take pictures, so they can be used for 3D reconstruction purposes. The game was designed to attract different types of players' personality. We archived this by integrating the use of different game play strategies, e.g. engaging in combats, collecting cards, exploring physical spaces, including social interactions, etc. The design of mechanisms to deal with sparse sensing and to evaluating data trustworthiness were among our biggest considerations.

In future work we have to compare the performance of our pure crowdsourcing application with this game-based strategy. We also have to investigate how to improve the security and reliability of the framework and how to handle intellectual property and privacy issues, e.g. reconstruction a person without the proper permission.

References

1. Simões, B., De Amicis, R.: Digital earth in a user-centric perspective. In: Fifth International Conference on Computing for Geospatial Research and Application (COM. Geo), pp. 47–48. IEEE (2014)

2. Michael, D.R., Chen. S.L.: Serious Games: Games That Educate, Train, and Inform. Muska & Lipman/Premier-Trade (2005)
3. Djaouti, D., Alvarez, J., Jessel, J.-P., Rampnoux, O.: Origins of serious games. In: Ma, M., Oikonomou, A., Jain, L.C. (eds.) Serious Games and Edutainment Applications, pp. 25–43. Springer, London (2011)
4. Marczewski, A.: Gamification: A Simple Introduction. Andrzej Marczewski (2012)
5. Zichermann, G., Linder, J.: The Gamification Revolution. Mcgraw-Hill Professional, Berkeley (2013)
6. Von Ahn, L., Dabbish, L.: Labeling images with a computer game. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 319–326. ACM (2004)
7. Von Ahn, L., Liu, R., Blum, M.: Peekaboom: a game for locating objects in images. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 55–64. ACM (2006)
8. Von Ahn, L., Ginosar, S., Kedia, M., Blum, M.: Improving image search with phetch. In: IEEE International Conference on Acoustics, Speech and Signal Processing, ICAAAP 2007, vol. 4, p. IV-1209. IEEE (2007)
9. Eickhoff, C., Harris, C.G., de Vries, A.P., Srinivasan, P.: Quality through flow, immersion: gamifying crowdsourced relevance assessments. In: Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 871–880. ACM (2012)
10. Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popović, Z., et al.: Predicting protein structures with a multiplayer online game. Nature **466**(7307), 756–760 (2010)
11. de Croon, G., Gerke, P.K., Sprinkhuizen-Kuyper, I.: Crowdsourcing as a methodology to obtain large and varied robotic data sets. In: 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014), pp. 1595–1600. IEEE (2014)
12. Tuite, K., Snavely, N., Hsiao, D.-Y., Tabing, N., Popovic, Z.: Photocity: training experts at large-scale image acquisition through a competitive game. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2011, pp. 1383–1392. ACM, New York (2011). http://doi.acm.org/10.1145/1978942.1979146
13. Bell, M., Reeves, S., Brown, B., Sherwood, S., MacMillan, D., Ferguson, J., Chalmers, M.: Eyespy: supporting navigation through play. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 123–132. ACM (2009)
14. Bartle, R.: Hearts, clubs, diamonds, spades: players who suit muds. J. MUD Res. **1**(1), 19 (1996)
15. Stewart, B.: Personality, play styles: a unified model. Gamasutra (2011). http://www.gamasutra.com/view/feature/6474/personality_and_play_styles_a_.php