Picture-based Localisation for Pervasive Gaming

Martin Fischbach, Jean-Luc Lugrin, Marc Erich Latoschik, Michael Fendt

Human-Computer Interaction
Universität Würzburg
Am Hubland
97074 Würzburg
E-Mail: martin.fischbach@uni-wuerzburg.de

Abstract: Localisation, i.e. determining the position of users or devices, constitutes a key requirement for almost all types of mobile pervasive games. However, current marker-based localisation systems, e.g. based on QR codes, present drawbacks that limit game deployment, scalability and maintainability. In this paper, we propose an alternative to solve these issues and introduce the first steps towards its full realisation. Our approach relies on markerless picture matching using a natural feature detection algorithm. Players reproduce a camera shot of a real-world site to confirm their presence and to progress further in the game. One of the critical requirements is to provide accurate recognition while preserving application responsiveness with a large range of mobile devices and camera resolutions. We developed a proof-of-concept system and determined the best picture resolutions and feature numbers necessary to preserve both accurracy and responsiveness on diverse mobile devices. Our first results demonstrate the feasibility to achieve precise recognition within realtime constraints. We believe such localisation system have the potential to considerably facilitate pervasive game authoring while promoting new types of game mechanics.

Keywords: Pervasive Games, Natural Feature Matching, Performance, Localisation

1 Introduction

Pervasive gaming describes a category of games that incorporate real-world activities and experiences in a virtual game world [BML05, KGB13]. Due to recent technological advances, pervasive games became increasingly feasible on affordable mobile hardware devices, like smartphones or tablets. Pervasive games allow the exploration of novel game concepts [KLL12] as well as the extension of traditional real or virtual games [GDB+07, GSZW+13, Inc14, Nia14]. Besides pure entertainment purposes, pervasive gaming also reveals great potential in application areas such as tourism [LGG13], edutainment [WW11, FGEM+11], or healthcare [SLB+10]. Localisation, i.e. determining the position of users or devices, constitutes a key requirement for almost all types of pervasive games [BOL+06]. Most pervasive games rely on players to be physically present in a specific geographical location to progress in the game. The required accuracy of the player's localisation may vary from one game to another. Rough geographic location estimation via GPS, mobile network cell, or WiFi hotspot information is widely used for pervasive games [BOL+06, KGB13]. When the player's location has to be estimated more precisely, artificial fiducial markers (e.g. QR codes) are usually employed [WW11, LGG13]. Although very robust, one major drawback of this approach is that the artificial markers have to be physically placed in the environment by the game authors. They oftentimes require to be frequently maintained to insure their presences and integrity. The placement of such markers is also not always possible or desirable as they may affect the site's aesthetics or offend against prescriptions. In addition, marker-based games have a low scalability due to their difficulty to be duplicated in various locations or to sustain long-term participation [NMW+12].

In this paper, we present an alternative to fiducial markers for player localisation based on natural feature tracking techniques. Such techniques are commonly used in pure augmented reality applications and provide a high percentage of recognition in realtime even with mobile devices [WRM+08, RRKB11]. Our approach relies on a picture-based localisation technique, where players have to confirm their presence at the right location, by reproducing a given picture with their mobile device. The content and perspective determined by the given photography is said to be a *Natural Reference Point* (NRP) that players should match as closely as possible. Besides the advantage of not requiring a physical marker to be placed in the environment, the creation of a NRP is also straightforward. The device's camera is simply positioned and oriented to specify the perspective the player has to discover and reproduce. The player is guided to the proximity of a NRP using global positioning data visualized via Google Maps. Ideally, other available sensor data, like the orientation obtained from a magnetometer, would be recorded to improve the matching process.

However, outdoor picture comparison using natural feature tracking brings a new set of challenges, especially if the support of a wide range of mobile devices is desired. High resolution pictures with numerous different distinguishable features are commonly required to be efficient. In addition, the utilized algorithms are typically computationally intensive and require a camera calibration process, preferably within an indoor environment. The pervasive game idea renders such ideal conditions impossible to achieve. An ideal solution has to be able to match pictures taken from different cameras and perspectives as well as content affected by environmental conditions (weather, season, and brightness). Moreover, its responsiveness has to support realtime feedback to guide players during their quest of matching the view point. Consequently, the identification of an algorithm capable of running on a multitude of devices with realtime constraints is a first step towards a fast and robust picture-based localisation for pervasive games. In the rest of this paper we will first briefly present related work, before introducing the overall concept and requirements of our approach. We then describe a proof-of-concept system and our empirical methodologies used to identify good candidate picture characteristics that will preserve both responsiveness and accurate recognition. Finally, we discuss the potential impact on pervasive game authoring and gameplay by outlining future improvements and a planned adaption of a commercial game product.

2 Related Work

The term *pervasive game* has been used to outline a heterogeneous set of game applications that "extend the gaming experience out into the real world" [BML05, KGB13] or "integrate the physical and social aspects of the real world" [MCMN05], ranging from smart toys through location aware, mobile, video or affective games to augmented reality and augmented tabletop games. Nieuwdorp [Nie07] carefully reviews existing definitions and emphasizes the equality of the terms *pervasive game* and *ubiquitous game*. In fact, she proposes the separation of two perspectives when denoting a game to be *pervasive*: a cultural perspective that describes the real-world gameplay and a technological perspective that describes the technologies supporting the game.

Technologically pervasive games depend on localisation, communication, context detection, augmentation, authoring and game engines, as well as orchestration and surveillance [BOL+06]. Although not all of these requirements are present or equally important amongst pervasive games, detecting the position of the user in the real world is essential in most cases. Besides the widely used global or local positioning systems, computer vision approaches provide another common solution for localisation or even tracking. Existent approaches are separated based on the requirement of an artificial fiducial marker. Marker-based approaches commonly utilize QR codes on mobile devices for localisation [WW11, LGG13], if the code's position is known prior, or for tracking in augmented reality applications [KTC09]. Markerless approaches build upon the detection of natural features in images, which are used to identify scenes or motifs from different perspectives. Compared to marker-based approaches, markerless approaches typically require more processing power [WS09]. Nevertheless augmented reality applications became increasingly suitable on mobile plattforms (e.g. [QCE14]), due to both recent advances in affordable mobile computing hardware and the development of highly efficient algorithms especially tailored for devices with limited hardware [WRM+08, RRKB11, PR11, PSR13]. These algorithms should also be suitable for picture-based localisation using natural features on mobile devices. Therefore, this paper presents first steps towards the full validation of the feasibility of this concept. We especially focus on the minimum picture resolution and number of matching features required to ensure an efficient comparison while preserving an enjoyable gaming experience.

3 Concept

Our main conceptual goal is to provide an alternative to QR codes, that can be utilized for identifying NRPs without the need to deploy an artificial fiducial marker in the environment. We propose to achieve this goal by applying natural feature matching to identify previously recorded reference points and thus confirm the user's proximity and attention. NRPs can be used in similar fashion to QR codes in pervasive games. For instance, to trigger the output of context information if the player successfully confirmed to be at a specific sight in a touristic game or to trigger the next step in the ongoing story line if the player has successfully confirmed to be at a designated remote location in a role-playing game. Besides the raw confirmation from the natural feature matching, other available sensor data can be used to support a player's task of finding and confirming a NRP (see figure 1).

Figure 1: Finding and confirming a NRP: (a) Global or local positioning systems are used to pilot the player into intermediate proximity. (b) Subsequently, a picture in combination with an optional description helps the player to identify the approached NRP. (c) Realtime feedback based on the device orientation obtained from a magnetometer and the quality of the natural feature matching guide the player to the NRP (concept). (d) If a certain threshold of the matching quality is exceeded, the NRP is considered to be confirmed.[1]

To implement the described interaction, the respective sensor data has to be recorded when defining new NRPs and has to be stored on a server. Since the required sensors already have to be present on the players' mobile devices, authoring is straightforward: in order to place or define a NRP, authors position their device like they want players to discover the NRP and confirm, similar to taking a photo. All necessary data can be recorded automatically without requiring authors to print a marker, deploy it in the environment and store its position on a server manually. Thus the usability of the authoring process is highly improved, promoting the integration of user generated content into pervasive game concepts by making the authoring a game itself. Altogether, the presented concept poses great benefits for pervasive games that rely on the players to discover specific locations or perspectives.

4 Proof-of-Concept

As a first step to prove this concept, a natural-featured-based picture comparison for the mobile operating system Android (Version 4.4) has been implemented. This prototype has subsequently been used to empirically determine picture characteristics that result in a good trade-off between accuracy and computing time.

[1] Stencils from https://www.graffletopia.com/stencils/578

4.1 Implementation

The picture comparison is built upon the open source C/C++ library *OpenCV* [Tea14] for image processing. The access to OpenCV from within the Android application is handled by the *Android NDK*. The comparison of two images is subdivided into three steps: (1) the detection of natural features in each picture, (2) the matching of the detected features between the pictures, and (3) the decision if two pictures are considered to contain the same scene or motif. The detection of natural features is accomplished by the application of the Oriented FAST and Rotated BRIEF (ORB) algorithm [RRKB11]. ORB is optimized for low-power devices, like smartphones, and already included in the OpenCV library. The feature matching strategy uses a *RANdom SAmple Consensus* (RANSAC) to assess matches [Lag11, Page 233ff.]. The final step is a binary decision that compares the number of matching features with a threshold.

4.2 Evaluation

The described picture comparison process depends upon several parameters, like image sizes or thresholds, that influence the overall performance and robustness. In order to determine optimal values for these parameters the utilized algorithms have been applied for two validation sets: V_1 containing 10 pictures of the same scene and V_2 containing 100 pictures of diverse motifs and scenes. To support the independence of the determined parameter values from a specific device, these pictures have been taken with different smartphones in different resolutions and formats. Figure 2 illustrates a sample comparison between two pictures of V_1.

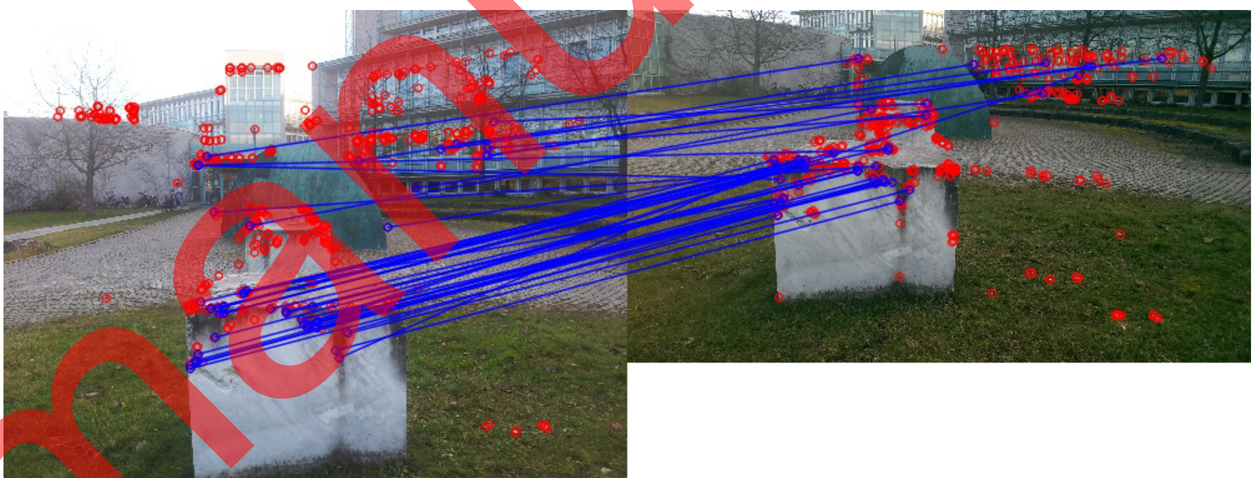

Figure 2: Sample comparison between two pictures taken by different devices. Red circles denote detected natural features. Blue lines connect pairs of matched features.

The first evaluated parameter in this study is the size of the images that should be passed to the picture comparison process to preserve both accuracy and responsiveness. The image resolution is positively correlated with the number of detected features and negatively correlated with the processing time. Therefore, the most critical aspect is to identify the lowest image resolution that will preserve a high fraction of detected features, necessary for accurate comparison. Figure 3 and 4 show the measured processing time and the number of detected features for the images of V_2. All images have been resized to several resolutions, by scaling the longer aspect to the desired length.

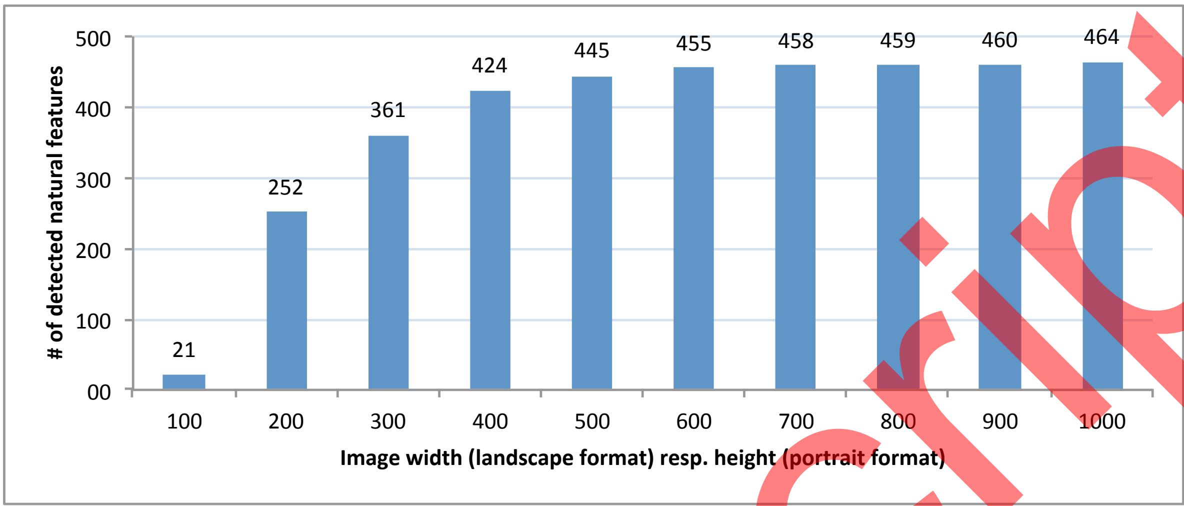

Figure 3: Average number of detected natural features for V_2. All 100 images have been scaled to all depicted resolutions.

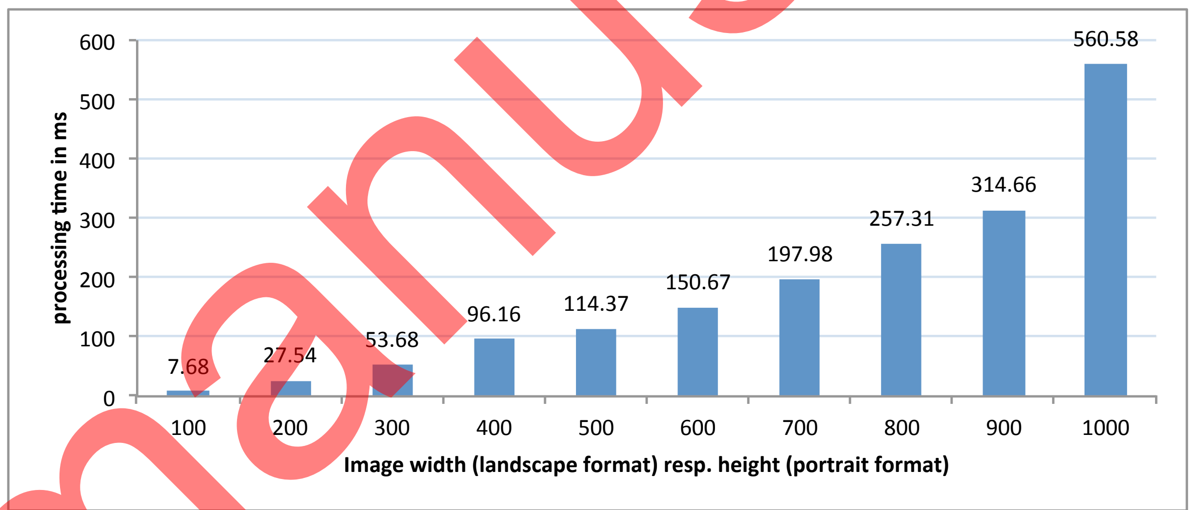

Figure 4: Average processing time for the detection of natural features. The measurements have been taken on a *Samsung Galaxy S2* using V_2. All 100 images have been scaled to all depicted resolutions.

Up to an image size of 500 pixels a high increase of the number of detected features can be noted. The increase between 500 and 600 pixels is only 2.3%, whereas the corresponding increase in processing time is 24.1%. An image size of 500 pixels is thus considered to be the best choice for further processing (see table 1). The required processing time of on average $114.37ms$ is only slightly above the limit of $100ms$, for system reactions that are perceived to be instantaneously [Nie94].

Device Name	Camera Resolution	Recommended Resolution
Samsung Galaxy S2	3264x2448 pixels	500x375 pixels
HTC One mini	2688x1520 pixels	500x283 pixels
HTC Wildfire S	2592x1728 pixels	500x333 pixels

Table 1: Recommended image resolutions for natural feature detection.

The second evaluated parameter is the threshold for the final binary decision. To get a lower bound for this threshold, every image from V_2 has been compared with all other images from this validation set, resulting in a total of 9900 direct comparisons (see figure 5). Because every image in V_2 has different content, the obtained quantities of matching features are assumed to be insufficient to denote a matching of two pictures. The highest obtained number of matching features is 44, excluding two outliers at 134 and 144.

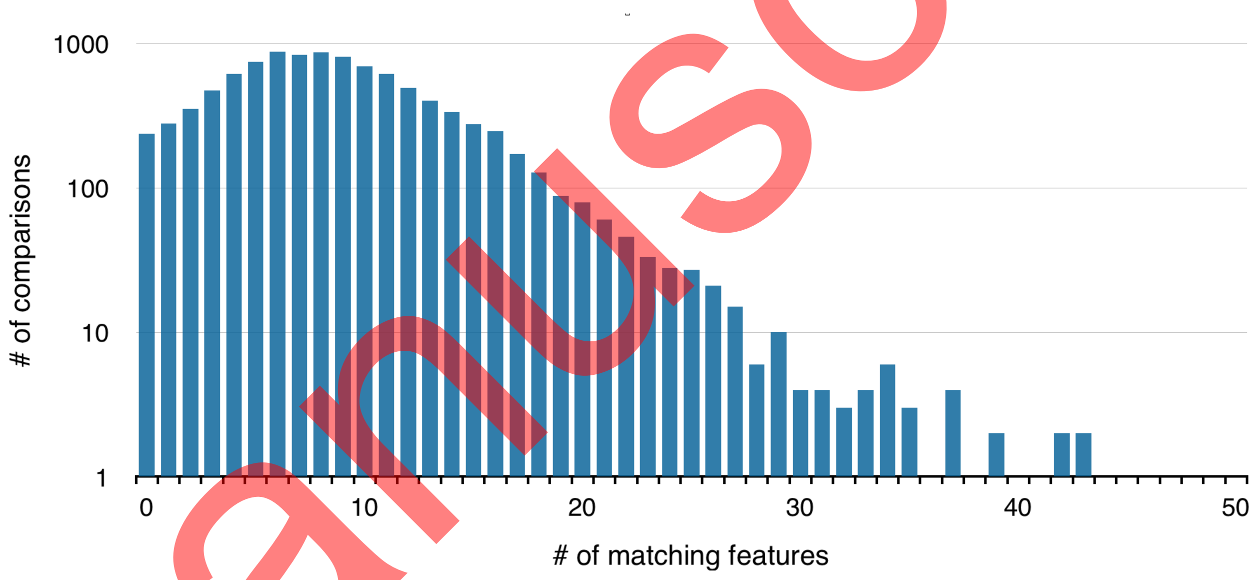

Figure 5: Number of matching features for all comparisons between pictures of V_2, showing diverse motifs and scenes (negative examples). Two outliers at 134 and 144 are not visualized.

To get an upper bound for the threshold, every image of V_1 has been compared in the same way, resulting in a total of 90 comparisons (see figure 6). Since every image in V_1 shows the same scene, the obtained quantities denote sufficient amounts of accepted matching features. The number of accepted matching features range from 12 to 191.

Because the determined bounds overlap, the positive and negative examples are not completely separable. While optimizing the binary decision by maximizing the corresponding F_1 score [MKS+99] would lead to a high accuracy, it does not meet the requirements of

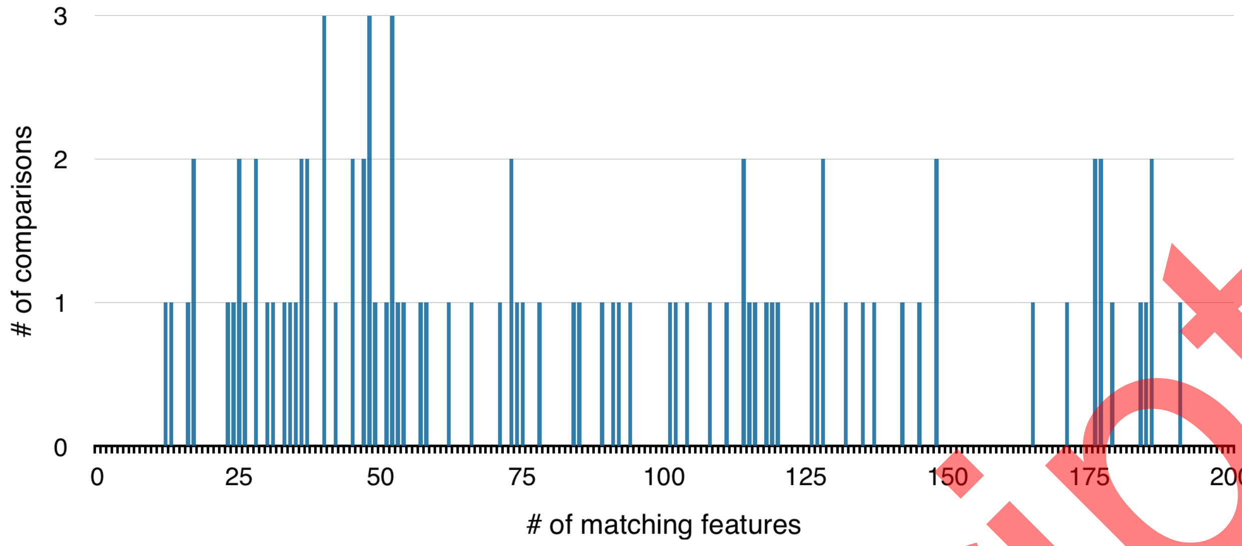

Figure 6: Number of matching features for all comparisons between pictures of V_1, showing the same scene (positive examples).

pervasive games. It is crucial not to mistake two pictures that do not show the same content to be equal (false positives). Therefore a threshold of 45, just above the highest occurrence of matches between two distinct pictures, is considered to be the best choice.

Even though this choice causes a higher amount of equal images to be denied (false negatives), the application context has the potential to neglect this flaw. Typically users will position their mobile device in the approximate location and orientation of the NRP, while slightly moving it. This allows the application to record several images that all can be checked against the corresponding image stored on the server. Resulting in a higher chance of confirming a NRP despite the presence of false negatives.

5 Conclusion

In this paper we proposed an alternative to improve the traditional deployement, scalability, and maintainability limitations of marker-based pervasive games. Our approach relies on markerless picture matching and uses the natural feature detection algorithm ORB, designed for mobile platforms. With our concept, players can localise themselves by reproducing a camera shot of a real-world sight in order to progress further into the game. One of the critical requirements was thus to provide accurate recognition, while preserving application responsiveness with a wide range of mobile devices and camera resolutions. We developed a proof-of-concept system and evaluated the best picture resolutions and number of features necessary to preserve both requirements on diverse mobile devices. Our first results demonstrate the feasibility to achieve precise recognition within realtime constraints.

We believe that many existing QR code-based pervasive games, e.g. [WW11, LGG13], could easily benefit from such an approach. It also has the potential to promote the development of novel gameplay concepts, especially for geocaching or role-playing games, where players are meant to discover specific perspectives of a historical or touristic site and the mounting of artificial markers is impossible. Finally, the simplicity of creating new NRPs improves the efficiency of the authoring process and thus could greatly facilitate the integration of user generated content into pervasive gaming concepts.

However, future work will evaluate to what extent the identified minimum resolution and number of features are supporting an enjoyable gaming experience, especially with a large diversity of images. The required degree of accuracy should promote a challenging gameplay without frustrating the player by making the game too difficult or easy. Consequently, the visualization of the degree of similarity as well as the orientation offset will be used to provide realtime feedback during the matching process. Before running a complete usability study with a full game implementation in partnership with a game company, we also wish to evaluate the system robustness to weather, season and brightness variations. The range of mobile devices under test will also be increased to a more representative sample, in order to continue the evaluation of our system's portability and performance.

References

[BML05] Steve Benford, Carsten Magerkurth, and Peter Ljungstrand. Bridging the Physical and Digital in Pervasive Gaming. *Commun. ACM*, 48(3):54–57, March 2005.

[BOL+06] Wolfgang Broll, Jan Ohlenburg, Irma Lindt, Iris Herbst, and Anne-Kathrin Braun. Meeting Technology Challenges of Pervasive Augmented Reality Games. In *Proceedings of 5th ACM SIGCOMM Workshop on Network and System Support for Games*, NetGames '06. ACM, 2006.

[FGEM+11] Faranak Fotouhi-Ghazvini, Rae A Earnshaw, Ali Moeini, David Robison, and Peter Excell. From E-Learning to M-Learning-the use of Mixed Reality Games as a new Educational Paradigm. *iJIM*, 5(2):17–25, 2011.

[GDB+07] Lyndsay Grant, Hans Daanen, Steve Benford, Alastair Hampshire, Adam Drozd, and Chris Greenhalgh. MobiMissions: the game of missions for mobile phones. In *ACM SIGGRAPH 2007 educators program*, page 12. ACM, 2007.

[GSZW+13] Anke Giebler-Schubert, Chris Zimmerer, Thomas Wedler, Martin Fischbach, and Marc Erich Latoschik. Ein digitales Tabletop-Rollenspiel für Mixed-Reality-Interaktionstechniken. In *Virtuelle und Erweiterte Realität, 10. Workshop der GI-Fachgruppe VR/AR*, Informatik, pages 181–184. Shaker Verlag, 2013.

[Inc14] Stray Boots Inc. Stary Boots Scavenger Hunts and Walking Tours. Retrieved from www.strayboots.com, 2014.

[KGB13] Vlasios Kasapakis, Damianos Gavalas, and Nikos Bubaris. Pervasive Games Research: A Design Aspects-based State of the Art Report. In *Proceedings of the 17th Panhellenic Conference on Informatics*, PCI '13, pages 152–157. ACM, 2013.

[KLL12] Ben Kirman, Conor Linehan, and Shaun Lawson. Blowtooth: A Provocative Pervasive Game for Smuggling Virtual Drugs Through Real Airport Security. *Personal Ubiquitous Comput.*, 16(6):767–775, August 2012.

[KTC09] Tai-Wei Kan, Chin-Hung Teng, and Wen-Shou Chou. Applying QR Code in Augmented Reality Applications. In *Proceedings of the 8th International Conference on Virtual Reality Continuum and Its Applications in Industry*, VRCAI '09, pages 253–257. ACM, 2009.

[Lag11] Robert Laganière. *OpenCV 2 Computer Vision Application Programming Cookbook: Over 50 Recipes to Master this Library of Programming Functions for Real-time Computer Vision*. Packt Publishing Ltd, 2011.

[LGG13] MaríaTeresa Linaza, Aitor Gutierrez, and Ander García. Pervasive Augmented Reality Games to Experience Tourism Destinations. In *Information and Communication Technologies in Tourism 2014*, pages 497–509. Springer International Publishing, 2013.

[MCMN05] Carsten Magerkurth, Adrian David Cheok, Regan L Mandryk, and Trond Nilsen. Pervasive Games: Bringing Computer Entertainment Back to the Real World. *Computers in Entertainment (CIE)*, 3(3):4–4, 2005.

[MKS+99] John Makhoul, Francis Kubala, Richard Schwartz, Ralph Weischedel, et al. Performance measures for information extraction. In *Proceedings of DARPA broadcast news workshop*, pages 249–252, 1999.

[Nia14] NianticLabs@Google. Ingress. The game. Retrieved from www.ingress.com, June 2014.

[Nie94] Jakob Nielsen. *Usability engineering*. Elsevier, 1994.

[Nie07] Eva Nieuwdorp. The Pervasive Discourse: An Analysis. *Comput. Entertain.*, 5(2), April 2007.

[NMW+12] Carman Neustaedter, Victoria Moulder, Ron Wakkary, Tejinder K Judge, and Anthony Tang. Designing Mixed Reality Games to Study Culture, Family Practices, and Social Engagement. 2012.

[PR11] Christian Pirchheim and Gerhard Reitmayr. Homography-based planar mapping and tracking for mobile phones. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*, pages 27–36, Oct 2011.

[PSR13] C. Pirchheim, D. Schmalstieg, and G. Reitmayr. Handling pure camera rotation in keyframe-based SLAM. In *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*, pages 229–238, Oct 2013.

[QCE14] Inc Qualcomm Connected Experiences. Vuforia - enable your apps to see. Retrieved from www.vuforia.com, July 2014.

[RRKB11] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. ORB: An efficient alternative to SIFT or SURF. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2564–2571, Nov 2011.

[SLB+10] Kevin G. Stanley, Ian Livingston, Alan Bandurka, Robert Kapiszka, and Regan L. Mandryk. PiNiZoRo: A GPS-based Exercise Game for Families. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology*, Futureplay '10, pages 243–246. ACM, 2010.

[Tea14] OpenCV Developers Team. OpenCV. Retrieved from http://opencv.org, June 2014.

[WRM+08] Daniel Wagner, Gerhard Reitmayr, Alessandro Mulloni, Tom Drummond, and Dieter Schmalstieg. Pose tracking from natural features on mobile phones. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 125–134. IEEE Computer Society, 2008.

[WS09] Daniel Wagner and Dieter Schmalstieg. History and future of tracking for mobile phone augmented reality. In *International Symposium on Ubiquitous Virtual Reality*, pages 7–10, 2009.

[WW11] Bian Wu and Alf Inge Wang. A pervasive game to know your city better. In *Proceedings of the 2011 IEEE International Games Innovation Conference*, IGIC '11, pages 117–120, 2011.