BUSINESS INTELLIGENCE THROUGH REAL-TIME TRACKING
Using a Location System Towards Behaviour Pattern Extraction

Pedro Abreu, Vasco Vinhas
FEUP - Faculty of Engineering of the University of Porto, R. Campo Alegre 1021, Porto, Portugal
DEI - Department of Informatics Engineering, R. Campo Alegre 1021, Porto, Portugal
LIACC - Artificial Intelligence and Computer Science Laboratory, R. Campo Alegre 1021, Porto, Portugal

Pedro Mendes
FEUP - Faculty of Engineering of the University of Porto, R. Campo Alegre 1021, Porto, Portugal
DEI - Department of Informatics Engineering, R. Campo Alegre 1021, Porto, Portugal
INESC Porto - Institute for Systems and Computer Engineering of Porto, Rua Dr. Roberto Frias, 378, Porto, Portugal

Keywords: Location System, Real-Time, Knowledge Extraction, Business Intelligence.

Abstract: Nowadays, tracking systems constitute an important knowledge support in order to compute important measurements in companys processes efficiency. As consequence of that, this project proposes a methodology and an application, based on a tracking system to obtain, by automatic means, dynamic location data on items. This solution assumes that the client carries or drives an item of some kind. In each item there is an identifying tag attached and hidden in order to make the item at hand detectable by all the sensors that are scattered around the area. Because of the fact that the tag is light and hidden and also has no information regarding the specific person/agent this process is completely transparent to the client or robot that is being implicitly tracked. This system produces real-time shop floor visualization maps with intelligible data on online item localization; individual item complete path routes; online and historical population density rates and path routes concentration; and also item vision enabled concentration maps as emulation for item omnidirectional vision considering occlusions. This proposed system might be useful in many different areas, for instance in a traditional retail environment tracing clients through a commercial area or enabling item tracking and route analysis in a hospital.

1 INTRODUCTION

Detecting behavioral patterns is a challenging task that marketing and distribution companies face. The issue has been addressed through the past years on several perspectives like deterministic psychology (Luce, 1999; Choustova, 2007). However due to their active consciousness, human beings are extremely unpredictable and so these methods failed to provide any accurate data that could be used for industrial purposes. Having these approaches been unsuccessful, statistical inference with large data sets is still one of the most powerful tools available.

Nowadays, tracking systems can represent a powerful tool to support monitoring activity. With these systems, some performance measurements on company process efficiency can be obtained regardless of the specific tracked target. This research work presents a methodology and a tool, based on a tracking system to obtain by automatic means, movements data on these elements.

The presented solution assumes that the client carries or drives an item of some kind, inside the space. It is also assumed that these items are outside the entry of the traceable space, although already inside of its admissible space. In each of the items there is an identifying tag attached and hidden in order to make the item at hand detectable by all the radars that are scattered around the area. By detecting it, the client or robot is being implicitly tracked in a completely transparent way for him/it since the tag is light and hidden and also has no information regarding the specific person/agent.

Several benefits can be withdrawn from using a system such as this. Instantaneously one could use it to monitor the traceable area in a more effective way

Abreu P., Vinhas V. and Mendes P. (2008).
BUSINESS INTELLIGENCE THROUGH REAL-TIME TRACKING - Using a Location System Towards Behaviour Pattern Extraction.
In *Proceedings of the International Conference on e-Business*, pages 51-57
DOI: 10.5220/0001909900510057

than looking at dozens of screens with images from security cameras. Live monitoring of the elements' positions on a specific area allows managers to identify congested sectors. It also allows him to identify hot and cold zones which may be related to the interests points at hand or, for instance, to a local security issue or other type of event such a medical emergence.

Regarding long term data analysis the advantages of using such a system are several and may be more than the ones presented next. By analyzing all the paths taken by the elements, it is possible to obtain the hot zones on any time frame and thus evaluate the success rate of a given promotional campaign and among other things. Erroneous and random movements may also be correlated with a security issue, and thus this system could also represent an interesting addition to conventional security systems. The results later obtained would clearly point out the success of the layout redesign. Of course these last measures imply correlating client positions with goods bought by him.

The paper is structured as follows: section 2 describes the current state of the art regarding the several areas of knowledge involved in the development of the system described in this paper. Section 3 describes the tool that simulates the environment as well as its architecture and core functionalities. Section 4 discusses the results obtained so far and the next section concludes the paper by summarizing the focus of this research work and pointing out future lines of discussion.

2 STATE-OF-THE-ART

Nowadays, tracking systems represent an important research area as their applications are transversal to several areas of knowledge (e.g computer science, medicine, simulation, robotics as well as industrial tracks). In the past few years, technology has evolved in order to provide more accurate measurements. In the robotics area, for better modeling the world, it is extremely relevant to accurately process the signals received by the multiple sensors involved. Locating objects of the real world to the modeled one is a critical task for the appliance of the navigation algorithms and methodologies. Following these advances the work published by Hyunwoong Park (Park H., 2006) presents a new kind of sensor system with multiple rotating range sensors. Such system allows a robot to guide itself on *a priori* unknown world. On the other hand these tracking systems also find interesting applications on scenarios where the context environment is already known. Regarding this last system, locating elements assumes a crucial role. To achieve this goal, several technologies have been used. By doing a brief comparison, it is observable that all of them have their strengths and flaws concerning characteristics like the cost in terms of initial investment and maintenance. There are others related to environment specificities. Among these last, other parameters such as coverable area, tracking detection errors and occlusion problems should also be considered.

One of the most effective technologies is also one of the most expensive ones and concerns detection of thermal signatures. This technique is appropriate to living organisms which emit particular heat waves. One particular application of this technology is the monitoring of the fauna in the ocean (Raizer, 2003). Another interesting technology is Bluetooth because most modern mobile equipments are prepared to send a receive data though this protocol. Although the initial investment is low the coverable area is not very wide and battery consumption is high, in relative terms (Jappinen P., 2007).

The cheapest solution is infrared based. Even though its price attractiveness, infrared systems tend to fail on most real environments because the signal is unable to reach the target if there is an opaque object between the receiver and the target (Krotosky J., 2007).

Two of the most emergent technologies for tracking are RFID and Wi-Fi based. The first one still lacks standardization which is somehow reflected in the pricing of both receivers and transmitters. It is based on high frequency radio waves having the detectable tags a passive or an active response. Passive tags are only detectable on a 13 meter radius and are used for instance on the new USA passports. Active tags, alternatively, are detectable on a much wider range but are more expensive (around 400%) essentially because the tags require an independent power supply (Chao C., 2007). Wi-Fi may also be considered as a tracking technology. This approach is mainly used for creating wireless computer networks but in this case the involved tracking only requires the usage of the low level protocols. This type of solution is interesting because it makes possible reusing existent computer networks for other proposes and takes advantages from possible simple detection with at least one access point. With only one access point the system's precision may not be very high but there is no need for triangulation. Occlusion problems and signals losses, with the use of this technique will be reduced to a residual level in both open spaces and indoors – considering that indoor spaces do not have significant metal structures within the walls) (Mingkhwan, 2006).

Another area that suffered several developments in the past few years concerns world modeling. In this research area there are relevant research topics, most of them related to computer graphics. Most of the current advances focus on three dimensional (3D) worlds. In this scope, the evolution on computer graphics is the most notorious. Nowadays, simple systems are able to represent complex 3D world including high resolution textures, detail animation (Vazquez, 2007) and weather condition (Grudzinski, 2007). It is even possible to recreate a 3D world from textual specifications (Moura J., 2004). On top of 3D world, many algorithms are applicable in other to optimize rendering performance and obtaining world data such as visible objects of a certain point.

These algorithms are too complex for most real-time tracking systems and therefore for this research work the world is assumed to be a two and a half dimension one (2,5D). In this scenario, a map is represented considering a *bird's eye view* and assuming that the height of the objects has no maximum value. In these conceptions it is simpler to obtain the set of potentially visible objects from a certain point using a portal culling algorithms; that consider walls as complete occluders and assigns a vision probability to each region in the map (Pires, 2001). Such is achieved by dividing the indoor space into separate sectors and then portals which represent the breaches between the sectors. By drawing cones that connect the observer point and both extremities of a given portal, one can obtain the areas where all the objects are potentially visible. Some caution is required when performing such an operation, since the lines that represent the vision cone cannot intersect the ones representing zone divisions. It is also relevant to state that one must assume the observer's vision direction is the center of a given "vision cone".

3 PROJECT DESCRIPTION

The project description is divided of three distinct subsections. In the first denominated as *Project Architecture* a description of the system's architecture is depicted. In the *Real Time Tracking Visualization and Concentration Maps* subsection the principles of the system features is explained and in the last section, *Client Vision Module* the used vision algorithm is exposed.

3.1 Project Architecture

This research work proposes a decentralized architecture and prototype tool following that same principle as detailed in Figure 1.

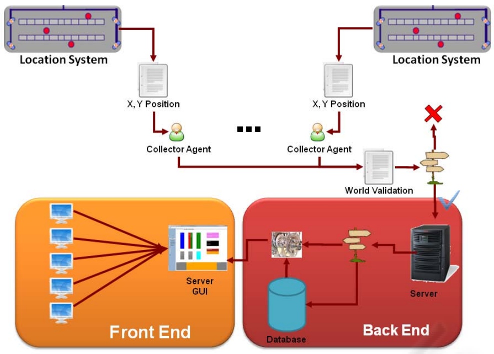

Figure 1: System's Architecture.

In the first stage, using a location system that might be instanciated in a RFID or Wi-Fi based solution, covering an area of for instance a large open space or building, with a maximum error of one meter, movement data is collected regarding the tagged elements present in the given specific floor. In order to gather the location data one agent has been developed to collect all the positions. This action has a given periodicity and is dependent on the location engine. Typically, this collected data is guaranteed to be obtained every second at most for every single tag, although this figure might be decreased depending on the number of simultaneously trackable objects.

After the collection process the agent sends the information to the server application. Prior to this action, the agent executes a simple, yet efficient, data validation that is based on the map of the structure that is sent by the server before the collector agent boot process.

The server application, before being able to receive any position data, must load the floor map that will contain the trackable objects, from a XML file. The maps are modeled as 2,5 D worlds and include several structures that can be easily adapted to many types of spaces. Before using it, the server validates the map against a XML Schema.

For instance, a XML file representing a traditional retail shop includes entry areas and an exit one that in this case is designated as a payment area. There are also walls that have infinitive height positioned around the map. In the supermarket example these last are named as shelves as can be seen on Figure 2. The proposed XML structure allows specifying a color and a designation for each half of the wall. The half is determined by the largest dimension, and in case of a square it is assumed to be vertically alligned as can be seen on Figure 2.

For each set of received coordinates, the server stores them in a database for universality sake. The

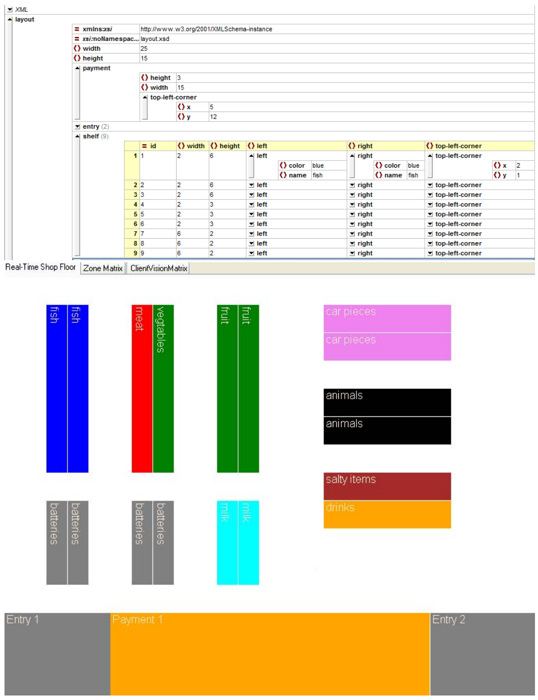

Figure 2: Drawn Map with Corresponding XML File Structure.

system executes it in a completely completely agnostic way concerning the database provider. At this point it is also relevant to state that the server is a multi-threaded application – having a thread per tag. Each thread writes into a reserved memory location the tag's current position. The previous position is not overwritten; instead it is stored in the database. Each thread includes some recovery proceedings such as deleting all the records of a specific tag in a given time frame if it stops transmitting its location for a long period of time regarding the context at hand.

As each thread executes, the server GUI is able to display online data. This data represents where the tags are located on the map. Other types of views involve processing the data in real-time, using computer memory and/or by consulting the database registers for further in depth reports (section 3.2).

3.2 Real-Time Tracking Visualization and Concentration Maps

The Server GUI includes several different views of both the online and historical data. In all these views there is a visual representation of the map. In order to draw a map the server requires to systematically perform a scale transformation involving real world coordinates and pixel coordinates. This kind of transformation must be dynamic because, in any instance, the GUI can be resized. When the view involves positioning tags on the map this transformation is also applied to their centre positioning.

The simplest view allows representing online the tags in their actual positions. Other views are obtained through the server's knowledgment extraction features. The Zone Matrix consists in determining in real-time which are the most and least populated zones, denominated as hot and cold zones.

It is also possible to consult this data on a wider time frame considering the same space with or without the same layout. This last feature requires database access. The zones are automatically obtained by dividing the space into a grid with flexible dynamic resolution. This dynamic division allows both a more in depth study of the hot and cold zones and also a less detailed one in order to study, for example, the spaces quadrants occupation.

Several other results are obtainable by accessing memory-based data structures for limited time frame analysis and by querying the database in similar modes. It is possible to obtain historical client paths, the shop areas walls, shelves or objects that were more observed by the clients.

In order to reproduce historical client path recognition all the clients coordinate are stored into the database with a timestamp that is related to a given map. Figure 3 exposes the paths taken by several clients in a given time frame.

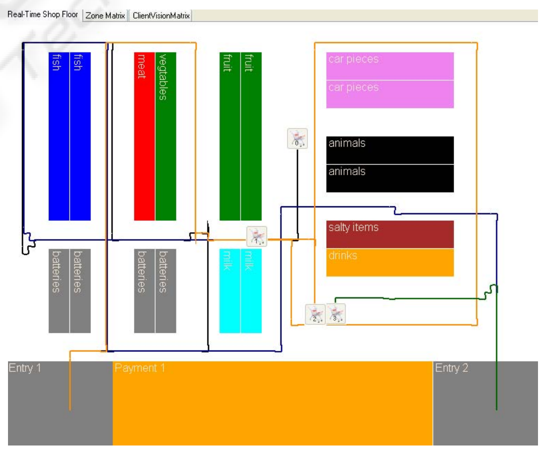

Figure 3: Paths taken by the Elements.

3.3 Client Vision Module

The client vision module uses a simplified version of the occlusion portal culling detection algorithm. This simplified version discards all zones that have a low probability of being seen and also disregards orientation and assumes that the observer can see simultaneously north, south, east and west directions. For each of these directions the observer throws a *vision cone* having its center coincident with the direction at hand.

The first intersected walls are considered to be visible, and the others invisible. The Figure 4 summarizes with descriptive colors which were the shelves that were most observed by the moving targets.

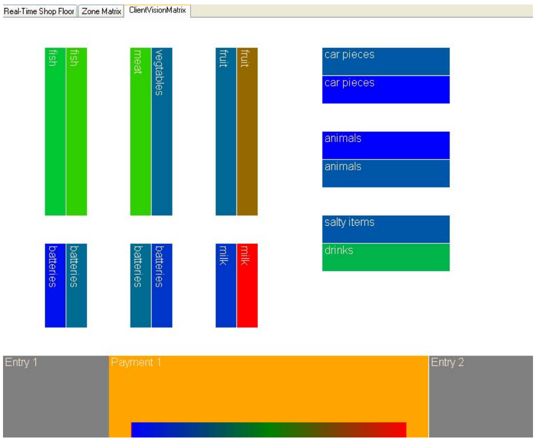

Figure 4: Visible Walls with Historical Data.

4 RESULTS

In this section, the project's results will be depicted taking into account three generic levels: system's main features; simulation statistics extraction and global aplication; and architecture stability and feasability.

For simulation purposes, there had been considered two standard computer configurations: a high-end machine with 4GB of RAM, an Intel Core 2 Duo E820 CPU and a SATA II 320 GB 16 MB cache hard disc denominated as configuration A, and a low-end configuration, denominated as B, equiped with a 2 GB of RAM, an Intel Pentium D 3.00GHz and a SATA II 250 GB 8MB cache hard disc. Both systems were equiped with Windows Vista Ultimate and the simulations were performed with similar workload conditions. In Figure 5 the experiment's results are fully depicted as for both hardware configurations four different scenarios were simulated. For each one it was recorded the CPU time needed to perform the most demanding task – real-time dynamic grid concentration levels with memory-based historical data – and the presence of absence of image flicker, with a different number of tracked items ranging from a single one to one thousand.

As previous note, one shall point that for single item tracking, the measured CPU time for both configurations is not available as the benchmarking tool reported zero seconds. The results showed that for ten items, the differences between low-end and high-end computers is absolutely negligible. For one hundred items, configuration B needs twice the time of configuration A but real-time visualization is not jeopperdized in anyway. In both cases, for this scale there was not registered any flicker effect and the process time was compatible with a real-time system. Only if the scale is pushed to one thousand, configuration B takes three point seven seconds to compute and even configuration A takes two point three seconds. These figures show that for this kind of scale it is needed a high-end computer system – even if one consider the traditional consumer market products – and hard real-time requirements are not met but one might still assume near real-time features that are perfectly adequated for this kind of management/monitoring systems.

			Configuration A	Configuration B
Number of Items	1	CPU Time (ms)	N/A	N/A
		Flicker	N	N
	10	CPU Time (ms)	15	23
		Flicker	N	N
	100	CPU Time (ms)	125	215
		Flicker	N	N
	1000	CPU Time (ms)	2300	3700
		Flicker	Y	Y

Figure 5: Simulation Performance Benchmarking.

Regarding the first aspect, all the enunciated predicted functionalities, thecnically described in the previuos section, were successfully implemented and fully tested. As illustrated in Figure 6, it is possible to visualise in real-time the location of up to one thousand items overlapped with the shop floor layout. This number of items can be increased but it is dependent of the external location system's features. In the same illustration, it is visible the tool's flexibility in what concerns to layout management and design as all shop floor static structures are fully defined and described through a simple, yet flexible XML configuration file. With this approach, it is possible to model heterogeneous environments and, therefore, apply the proposed system to several domains.

In spite of the importance of the mentioned functionalities, the greated added value resides in the knowledge extraction extendend features. Having in consideration the online item location gathering, the system is able of real-time item path reconstruction and visualization operating both in memmory-based or database access, depending on data dimension. Conducted experiments showed that real-time memmory access is feasible using a low-end computer – with 2 GB of RAM – for tracking one hundred items for a period of an hour at a medium pace.

Figure 6: Real-Time Item Location.

As illustrated in Figure 7 the location data is used to extract more significant information about item dispersion/concentration both in the present and also considering historical data. It was used a gradient scale where concentration levels vary through the RGB scale where red means high levels of concentration and blue low levels.

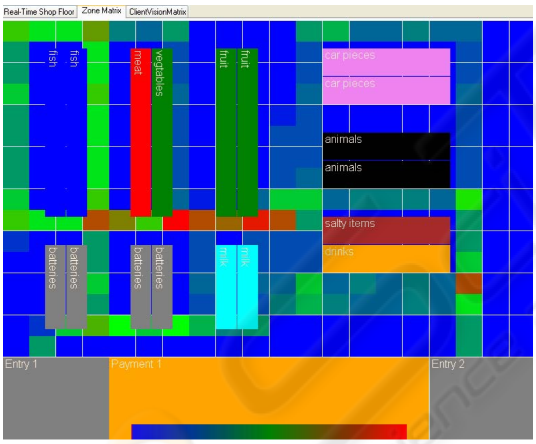

Figure 7: Historical Concentration Matrix.

One significant functionality of this model is the possibility to perform calculations based on completely flexible and dynamic projection grid. This option proved to be efficient on online data processing for a significant number of tracked items - approximately one hundred - without database access by using a temporal location matrix. This feature enables a full detailed concentration analysis in real time when recurring to a high definition grid that divides the layout in small areas; and enables swift high *big-picture* studies when using a less tight net. This capability is appropriately described in Figure 8.

Finally, considering the features results description, one shall paint the relevance of the vision module.

Figure 8: Dynamic Concentration Grid Example.

This application requirement performs the emulation of an omnidirectional vision of each trackable item. The described algorithm is able of identifying the visible objects, having into consideration both single instantaneous data and historical information, previously collected and stored – in direct memmory access or in a database. In the conducted experiments, this system's module also showed high levels of efficiency and correctness; much similar to the ones already described in the above paragraph.

Regarding the system's global architecture definition and implementation, the undertaken simulations demonstrated its adaptative capability through its flexibility in what concerns to both the database provider and, perhaps more important, to the location system tecnology. These caracteristics greatly enhance the whole system's applicability in several scenarios. Still in this domain, the distributed system's design enables the usage of low-end computers. Therefore it constitutes an incentive to client's IT infrastructure reusage while minimizing the solution's economical impact. Simultaneously, this approach enables greater site manager's empowerment through real-time information access to all system's features visualization. These actions can be triggered for both partial and global organization providing more and deeper analysis points of view.

5 CONCLUSIONS

Considering the project's simulation environment and the achieved results, one shall state that, although the location engine had been implemented in order to realistically simulate traceable items, all concept has been demonstrated. The developed prototype proved to be efficient and effective in large scale distributed data gathering and real-time item location visualization.

Taking into account the system's architecture, it was verified the concern in allowing multi-store management with both distribuited modules and central integration concerns that enable consistent and online knowledge extraction and visualization. Having in mind the different application modules, one ought to refer that the integration with the XML-based layout manager proved to be extremely flexible to accommodate distinct real scenarios and yet realizable and realistic.

Considering the most noble project's slice, one ought refer to the previously enunciated knowledge extraction features. Having as support the results depicted in the last section, it is secure to state that the system is able to produce real-time shop-floor visualization maps with intelligible data on online item localization; individual item complete path routes; online and historical population density rates and path routes concentration; and also item-vision enabled concentration maps – as emulation for item omnidirectional vision, yet considering occlusions. All of these features are allowed for graphical user interface through different grid dimensions for distinct analysis granularity. Bearing in mind the project as whole, one shall state that the developed knowledge extraction platform with online and diversified visualization tools constitutes a solid ground for online item tracking and heterogeneous space management with distributed capabilities. One final major advantage of the proposed system resides in the total independence regarding the external position engine both in terms of suppliers and even more important in terms of base technology.

In spite of the enunciated project's accomplishment, an even by being in prototype stage – yet reliable and fully functional – there are several future work areas that are able to greatly enhance the system's appliance and success. From these, the most relevant ones are believed to be centered in eccentric shop-floor layouts both in terms of shape and multi-level buildings; complete path routes analysis enabling common node fusion for global paths probabilistic construction; flexible and dynamic report definition tool with configurable alarm triggering; and perhaps the most interesting would be the characterization of 'what-if' scenarios with simulated traffic based on real historical data. Considering both the project current achievements and the depicted future work areas, one might identify the most desirable implementation domains. Although there are not limited to, the proposed system might be useful for traditional retail environment for shopping cart tracking; tracing clients through a commercial area such as shopping centers; enabling item tracking and route analysis in an hospital; producing activity reports and analysis in controlled areas such as penal complexes, mental institutions or closed educational organizations.

ACKNOWLEDGEMENTS

The authors would like to thank Professor Eugnio Oliveira for inspiration and the contribution regarding behavior pattern recognition and artificial intelligence guidelines, Professor Augusto Sousa for enlightenment in the portal cells algorithms.

REFERENCES

Chao C., Yang J., J. W. (2007). Determining technology trends and forecasts of rfid by a historical review and bibliometric analysis from 1991 to 2005. In *Technovation 27, Issue 5*, pages 268–279. Elvisier Ltd.

Choustova, O. (2007). Price-dynamics of shares and bohmian mechanics: Deterministic or stochastic model? In *Foundations of Probability and Physics*, pages 274–282.

Grudzinski, J. (2007). Real-time clouds and weather simulation in computer games. In *Learning with Games 2007*.

Jappinen P., P. J. (2007). Preference-aware ubiquitous advertisement screen. In *eCommerce 2007 Proceedings*, pages 99–105. IADIS Press.

Krotosky J., T. M. (2007). A comparison of color and infrared stereo approaches to pedestrian detection. In *Intelligent Vehicles Symposium*, pages 1226–1231.

Luce, R. (1999). Where is mathematical modeling in psychology headed? In *Theory and Psychology*, pages 723–737.

Mingkhwan, A. (2006). Wi-fi tracker: An organization wi-fi tracking system. In *IEEE CCECE/CCGEI, Ottawa,*, pages 1353–1356.

Moura J., Coelho A., L. P. e. a. (2004). *Readings in Information Society*. Universidade Fernando Pessoa.

Park H., Lee S., C. W. (2006). Obstacle detection and feature extraction using 2.5d range sensor system. In *SICE-ICA SE, 2006. International Joint Conference*, pages 2000–2004.

Pires, P. (2001). Dynamic algorithm binding for virtual walkthroughs. Master's thesis, Technical University of Lisbon.

Raizer, V. (2003). Validation of two-dimensional ocean microwave signatures. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03.*, pages 2694–2696.

Vazquez, P. (2007). Real time falling leaves. In *Second Internacional Conference on Computer Graphics Theory and Applications*, pages 244–251.