Check for updates

Hindawi Publishing Corporation
Advances in Mechanical Engineering
Volume 2013, Article ID 234747, 10 pages
http://dx.doi.org/10.1155/2013/234747

Review Article

Autonomous Vehicle Navigation Using Vision and Mapless Strategies: A Survey

Mehmet Serdar Güzel

Computer Engineering Department, Ankara University, 06560 Ankara, Turkey

Correspondence should be addressed to Mehmet Serdar Güzel; mguzel@ankara.edu.tr

Received 11 February 2013; Accepted 25 July 2013

Academic Editor: Shao Zili

Copyright © 2013 Mehmet Serdar Güzel. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This survey addresses the existing state of knowledge related to vision-based mobile robots, especially including their background and history, current trends, and mapless navigation. This paper not only discusses studies relevant to vision-based mobile robot systems but also critically evaluates the methodologies which have been developed and that directly affect such systems.

1. Introduction

A robot is defined as a programmable, self-controlled device. It is in essence a machine which is able to function in place of a living agent [1]. Mobile robots have a long history. Shakey, the world's first mobile robot, was developed in the late 1960s at SRI's Artificial Intelligence Centre (Stanford Research Institute) [1]. Not surprisingly, it has had a substantial influence on present day robotics. Shakey was equipped with various sensors and driven by a problem-solving program called "STRIPS," and used algorithms for perception, world modelling, and actuation. Low-level action procedures were responsible for moving, turning, and path planning tasks. The high-level program could make and execute plans to achieve goals. Another example of an early robot is CART which was developed at Stanford University in 1977 by Hans Moravec as part of his doctoral thesis [2]. However, CART was very slow, not because it was slow-moving by design, but because it was "slow-thinking."

The main reason for this was the difficulty of processing vision data using slow computer processors [3]. Another well-known example was Rover developed at Carnegie Mellon University (CMU) in the early 1980s. Rover was equipped with a camera and was able to perform better than CART. Nevertheless, its thinking and acting were still very slow [4]. In the last decade, the main developments in the area of robotics have come through technological breakthroughs in the areas of computing telecommunications, software, and electronic devices. These technologies have facilitated improvements in intelligent sensors, actuators, and planning and decision-making units which have significantly increased the capabilities of mobile vehicles. The latest trend in robotic intelligence is toward imitating life, for instance in evolutionary robots and emotional control robots.

Another area of technological challenge for the next decade is the development of microrobots and nanorobots for medical applications. On top of this, a paramount challenge will be to find an appropriate balance between human-assisted systems and fully autonomous systems and to integrate technological capabilities with social expectations and requirements [1].

Vision-based indoor mobile robot navigation has been studied for decades and is one of the most powerful and popular sensing method used for autonomous navigation. Compared with other on-board sensing techniques, vision-based approaches to navigation continue to demand attention from the mobile robot research community, due to their ability to provide detailed information about the environment which may not be available using combinations of other types of sensors. The great strides achieved in the area of vision-based navigation systems are significant; however, there is still a long way to go. Two important survey papers that have been published review various aspects of the progress made so far in vision for mobile robot navigation [5, 6].

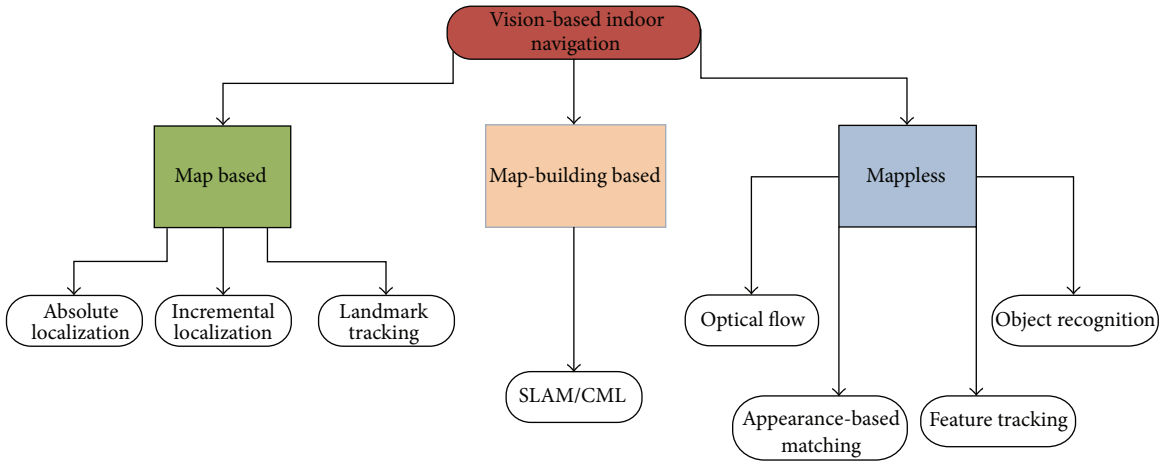


Figure 1: Vision-based indoor mobile robot navigation techniques.

This paper mainly discusses their classification of vision-based navigation systems and attempts to reveal the appropriate state-of-the-art for the indoor environments. A corresponding schema, summarizing state-of-the-art of vision-based navigation, is illustrated in Figure 1. According to the given figure (see Figure 1), there two methods are developed based on map, namely: map-based and map-building-based approaches which will be detailed in the following section.

2. Vision-Based Navigation Using Map

This section aims to detail techniques employing or constructing map to complete navigation problem especially for indoor environments. The first part of this section will introduce map-based navigation approaches in which the robot requires either a metrical or topological map to achieve the given navigation tasks. The second part will highlight the map-building-based methodologies which construct and adapt maps to navigation problems simultaneously. To understand these concepts, a simple analogy was described in [7], and a similar analogy including a daily life scenario is described next. In the scenario it is assumed that a visiting researcher is in Newcastle city centre and needs to return to the Stephenson Building (Newcastle University, Claremont Road) where his office is located. There are various methods that he can follow to reach the office. First, he could memorize the number of steps walked from the university, and he could return by counting the same number of steps. This would be dead-reckoning navigation. He could also buy a map of the environment in order to reach the goal, as in map-based navigation. However, this solution entails that somebody has previously named the streets and have drawn the map. Alternatively, he could also draw his own map using map-building navigation while exploring the city, but this would cost a lot of time and effort. Finally, he could look around trying to find the Claremont Tower and then try to approach it keeping the top of the tower in his field of view (map-less navigation). The goal is to reach the tower, since he knows that Daysh Building is next to it. Autonomous navigation architectures utilize some of these solutions to track a trajectory towards the required goal. Dead-reckoning navigation is the cheapest method and essentially includes an odometry system. However, this solution may include many mechanical problems that produce an increasing error which is unacceptable in long-term navigation. So, an additional perception system is mandatory. Vision is perhaps the most broadly researched perception system.

2.1. Map-Based Navigation. Many techniques employ metric or topological maps to navigate. Navigation techniques need certain knowledge of the environment, and maps may contain different degrees of detail, varying from a complete CAD model of the environment to a simple graph of interconnections between the elements in the environment. One of the key classification criteria in this approach depends on the type of map. For instance, metric-based maps generally favour techniques which produce an optimum, while qualitative methods such as topological maps employ identifiable gateways or landmarks to produce a route.

The main idea behind map-based navigation is essential to provide the robot with a sequence of landmarks expected to be found during navigation, and the task of the vision system is then to search for and recognize the landmarks observed in an image. When the landmarks are recognized, the robot can employ the map to estimate its own position (self-localization) by matching the observation (image) against the expectation (landmark) description in the database. The steps of vision-based localization can be divided into four steps [5].

Acquiring Sensory Information. Acquiring images with associated devices.

Detect Landmarks. Extracting edges, smoothing, filtering, and segmenting regions.

Matching. Identifying landmarks by searching possible matches using the database based on predefined conditions.

Calculate Position. Whenever a set of matches is obtained, the system needs to calculate the position using the observed landmarks' positions in the database.

The start position of the robot is unknown under absolute localization methods. Accordingly, the system must provide exact matching between the current and expected data, derived purely from the entire database. This self-localization problem has been solved either using deterministic triangulation or Monte Carlo-type localization [5, 6]. A detailed implementation of the Monte Carlo localization method to localize a mobile robot without knowledge of its starting location was proposed [8]. Incremental localization assumes that, at the beginning of the navigation, the position of the robot is known approximately. In such cases, the localization algorithm basically keeps track of uncertainties in the robot's position as it executes motion commands and, when the uncertainties exceed a limit uses its sensors for a new fix on its position [5]. The FINALE system is a good example of being able to achieve incremental localization using a geometrical representation of space and a mathematical model of uncertainty regarding the location of the robot [9]. Stereo vision is also preferred by researchers in order to reduce errors [10]. Christensen et al. [11] employed CAD models to represent the working space combined with stereo vision to reduce errors [11].

The final method is landmark tracking in which landmark tracking algorithms determine the position of the robot, detect landmarks on the camera image, and track them in the successive scenes. Landmarks can be artificial or natural. In both cases, the robot needs to recognize the landmarks in order to be able to track them [12–14]. Kabuka and Arenas [12] was the first employing artificial landmark tracking for vision-based navigation. A good example of natural landmark tracking-based navigation was proposed in [13] in which a landmark detection technique that assures stable detection even under variable brightness and an obstacle measurement technique that combines obstacle region segmentation with stereo vision were proposed.

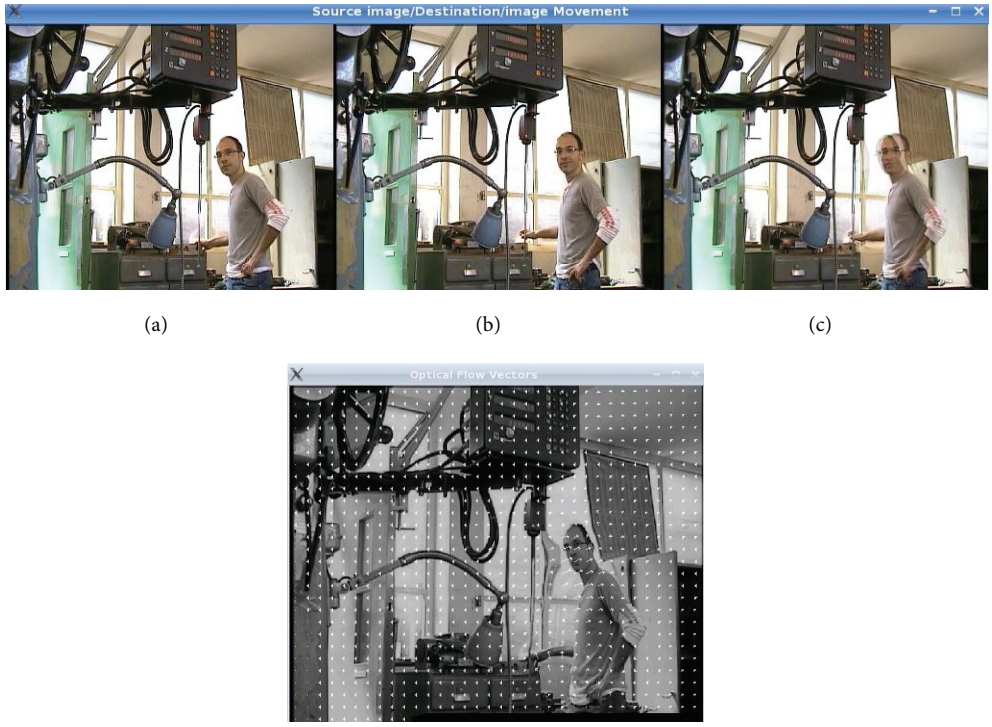
2.2. Map-Building Based Navigation. Sometimes modelling an environment could be difficult particularly if one also has to provide metrical information. An alternative navigation strategy, the map-building-based approach, has been used in both autonomous and semiautonomous systems that entails searching the environment and building a representation of it.

One of the earliest attempts at a map-building technique was carried out by the Stanford CART Robot equipped with a camera. Subsequently, an interest operator algorithm was improved to detect 3D coordinates of the images [15]. The system basically demonstrated the 3D coordinates of the objects, which were stored on a grid having 2 m cells. The map was updated at each iteration; and obstacles were also shown in the map. But the most important problem with the whole system was performance, which took five hours to go 20 metres. Visual navigation studies, employing map-building-based strategies, from the late 1990s to the present, have focused on two methodologies, namely: simultaneous localization and mapping (SLAM) or concurrent mapping and localization (CML) [6]. These methodologies principally propose solutions to automatically overcome the problem of the exploration and mapping of any unknown environment, which essentially entails three simultaneous tasks comprising navigation, mapping, and localization. Vision-based SLAM/CML algorithms mainly employ stereo vision as primary sensor. Se et al. implemented a vision-based mobile robot localization and mapping system in which the robot was equipped with a stereo system to build a 3D map so as to localize simultaneously in 3D [16]. The map was represented as a scale invariant feature transform (SIFT) feature database. It was constantly updated frame by frame and was adaptive to dynamic environments. An alternative and efficient solution to the SLAM problem based on a pair of stereo cameras has also been recently proposed which employs 3D landmarks to localize the robot, as well as constructing an occupancy grid for safe navigation [17]. Other map-building-based navigation techniques are those that impose a human-guided training stage. In such solutions, a human operator guides the robot through an unknown environment. During this process, the robot records images with a stereo camera and constructs the 3D map incrementally. After the map is built, the robot tracks extracted features and computes the optimum path [18].

3. Mapless Navigation

This section discusses a representative selection of mainly reactive visual-based navigation techniques in which navigation is performed without any prior description of the environment. Mapless navigation is scarcely new compared with the previously defined solutions, but new projects using vision systems have been developed in several directions in the last few years. In the systems surveyed in this section, no maps are ever created. The robots can navigate by observing and extracting relevant information about the landmarks in the environment. These elements can be objects such as desks, boxes, and doorways. The mapless visual navigation techniques discussed here are classified in accordance with the main vision technique or types of clues used during the navigation which are optical flow, appearance based, and object recognition navigation techniques based on feature tracking (see Figure 1) [5].

3.1. Optical Flow for Robot Navigation. Optical flow is defined as the motion of all the surface elements from the visual world. When a person moves through the world, the objects in the visual environment flow around this person. The human visual system can detect that person direction of travel from the movement of these objects. Optical flow can be defined as the apparent motion of features in a sequence of images, as shown in Figure 2. It is believed that when the insect is in relative motion with respect to the environment. Accuracy and the range of operation can be altered by changing the relative speed. For instance, features such as "time-to-contact" (depending on speed) are more relevant than distance when it is necessary to avoid an obstacle [5]. Bernardino and Santos-Victor developed an optical flow-based navigation system imitating the visual behavior of bees, called *robee*, and was equipped with a stereo vision system, mimicking the centring reflex behaviour used by a bee



(a) (b) (c)

(d)

Figure 2: Motion estimation, (a) source image, (b) destination image, (c) motion estimation, and (d) estimated flow vectors.

to navigate safely [19]. The robot localizes itself using the difference between the velocity of the image seen with the left eye and the velocity seen in the right eye. If the difference is close to the zero, the robot keeps moving forward. However, if the velocities are different, the robot moves toward the side whose image changes with a lower velocity. Several successful navigation systems have been recently inspired by this centring reflex and implemented to navigate a mobile robot through an unknown indoor environment. For instance, a mobile robot platform utilizing a binocular vision system to estimate optical flow in some way emulates corridor following behaviour to navigate [20]. Duchon et al. implemented a monocular vision-based navigation system based on optical flow algorithms and action modes (behaviours). Simulation and real experiments revealed that the system was capable of navigating in a maze whilst successfully avoiding obstacles [21]. Since then, several optical flow-based control algorithms integrated into a behavioural model have been implemented to be used in real-time applications [22–25].

Additionally, a simple but efficient optical flow-based visual-based navigation system was developed in [26] under the DARPA-Mobile Autonomous Robot Software (MARS) program. The proposed technique first estimates image edge maps by detecting Laplacian of Gaussian (LOG) zero crossings. Afterwards, the edge maps of consecutive frames are used by a patch matching procedure to calculate corresponding optical flow.

3.2. Appearance-Based Methods. Appearance-based methods fundamentally rely on the idea of memorizing the working environment. The main idea is to store images or templates of the environment and associate these images with commands that will steer the robot to its final destination [5]. These methods mainly consist of two procedures. The first one is the training phase in which images or prominent features of the environment are stored as model templates. The models are labeled with certain localization information, which are then associated with appropriate steering command. Secondly, in the navigation stage, the robot has to identify the environment and localize itself by matching the current image with the stored templates [6]. The main problems with the method are to find an appropriate algorithm for the representation of the environment and to define the on-line matching criteria [5, 6].

Beginning in the late 90s, a well-known example of this method was proposed in which the robot merges visual information and their azimuths to build up a representation of its location which is used to estimate the best movement to reach the goal. Another popular technique derived from this approach is the *view-sequenced route* representation which primarily focuses on route construction using a sequence of images and a template matching algorithm to guide robot navigation. Captured images are used to form a sequence of images. Each image in the sequence is associated with the motions required to move to a corresponding destination.

This approach basically introduces the concept of visual memory [27]. Multidimensional histograms provided by the statistical analysis of images are an alternative method to guide mobile robots in appearance-based strategies. Statistical data, including that related to colour, edge density, and texture, are utilized to build a multidimensional histogram database. The recognition of the environment during the navigation stage is achieved by matching the multidimensional histogram of the current image with the multidimensional histogram of the stored templates. This technique consumes less computational resources than when using correlation algorithms [28]. Recent groundbreaking research has proposed an entirely qualitative method in which feature points are automatically detected and tracked throughout the image sequence. In the teaching phase, the (KLT) feature tracker computes the displacement and minimizes the sum of the squared differences between consecutive image frames. The feature coordinates in the replay phase are compared with those computed previously in the teaching phase in order to estimate the steering commands for the robot. Experimental results revealed that the capability of autonomous navigation in both indoor and outdoor environments was successful with the proposed method [29].

An important concept in visual-based mobile robot navigation is the idea of visual homing, inspired by insect behaviour. Insects are able to return to important places in their environment by storing an image of the surroundings while at the goal, and later computing a home direction from a match between this snapshot image and the currently perceived image. For instance, an agent employing a visual homing algorithm captures an image at the home location. It then attempts to return to this location from a nearby position. It compares the current image with the snapshot and infers the direction to the location of the goal from the disparity between these images. It is considered that these aspects of insect behaviour can be a basis for the development of robust navigation algorithms for mobile robots [30].

Visual homing is an appearance-based navigation strategy whose homing algorithms are based on image-based holistic methods using disparities between whole images to compute homing vectors. Image warping is a popular method which is considered to be one of the most reliable visual homing methods for indoor use in this category. It involves calculating the set of all changes in position and orientation between the current and snapshot images. Warping methods distort the current image as if the agent would move according to certain movement parameters. The space of possible movement parameters is then searched for the parameter combination leading to the warped image that is as similar as possible to the stored image. In order to achieve this, warped images are compared to stored image using a pixel-by-pixel correlation measure. The current home vectors are determined based on the strongest similarity between those images [31]. A new and simple visual homing algorithm has been recently proposed where the home position is considered as a charging station [32]. The algorithm basically utilizes the root mean square (RMS) difference and exclusive or (XOR) functions to compare current and snapshot images. The paper reveals that it is possible to implement homing algorithms which allow a robot, fitted with a panoramic camera, return to a reference position from any starting point in an area [32].

3.3. Object Recognition. For the appearance-based approaches previously mentioned, the robot is only able to access few sequences of images that help it to reach its final destination or use predefined images of target goals that it can use to track and pursue. An alternative method has been proposed which essentially employs a symbolic navigation approach instead of memorizing the environment [33, 34].

In this case, the robot utilizes symbolic commands such as "go to the corner in front of you" or "go to the main exit." For instance, a command such as "go to the corner in front" informs the robot that the landmark is a corner and the path points straight ahead. The robot builds a map called an "s-map" which is a 2D grid that stores the projections of the observed landmarks as they are recognized. After the location of the landmark is projected into the s-map, the robot generates the path using GPS-like path planner and employs odometry to reach the goal [5].

3.4. Navigation Techniques Based on Feature Tracking. Tracking the motion of moving elements, including lines, corners, or specific regions in a video sequence has become a popular and robust way of performing navigation. Techniques inspired from tracking are called feature-based approaches which, in essence, determine the trajectory and motion of the robot by tracking and finding relative changes in the position of extracted features [6].

This category can also include feature-based visual homing strategies. Feature-based methods fundamentally segment snapshot and current images into landmarks and background. To operate successfully, feature-based navigation algorithms must extract the same features from snapshot and current images (the feature-extraction problem). Each feature extracted from the *snapshot* image must then be paired with a feature from the *current image* (the correspondence problem). The feature extraction and correspondence problems are difficult to solve in cluttered environments in real-time, since the appearance of landmark changes with viewpoint [31].

One of the earliest studies regarding feature tracking systems introduced a fruit tracking system employing the size and position of a valid fruit's regions in colour images, to control the motion of a fruit-picking robot [35]. Trahanias et al. implemented a robotic system that is able to extract landmarks automatically in indoor environments, using a selective search for landmark patterns which relies both on the workspace and the distinctiveness of the objects in the environment [36]. A KLT tracker-based homing schema was rooted in the extraction of very low-level sensory information, namely, the bearing angles of corners. This was implemented on a robotic platform to evaluate the results [37]. Alternatively, in most cases, feature tracking-based navigation algorithms do not provide an obstacle avoidance module, which must therefore be implemented by other

means depending on the problem. For instance, Li and Yang proposed a behavioral navigation architecture for mobile robots, which utilized a robust visual landmark-recognition system based on genetic algorithms to guide the robot, which used a fuzzy-based obstacle avoidance system and ultrasonic range finder [38].

Two images of the same planar surface in space are related by a homographies. This concept has many practical applications, including mobile robot navigation. In a recent study, Guerrero et al. introduced a method based on homographies computed between current images and images taken in a previous teaching phase with a monocular vision system. The vertical lines (features) were used to estimate the homographies, which are automatically extracted and matched. A complete homography motion can compute rotation and translation up to a scale factor [7]. Besides, a simple bearing only measurement technique relying on monocular vision is proposed with odometry readings in which the reference image and the current image are compared to enhance the localization accuracy of the robot [39]. The scale invariant feature transform (SIFT) method, detailed in Figure 3, is a milestone among techniques to detect the features of images or relevant points, and nowadays it has become a method commonly used in landmark detection applications [40, 41]. The SIFT extracts features that are invariant to image scaling, rotation, and illumination. During the robot navigation process, invariant features which have been detected are then observed from different points of view, angles, and distances and under different illumination conditions and thus become highly appropriate landmarks to be tracked for navigation.

Pons et al. employed the SIFT algorithm for feature-based homing, and are utilized to recover the misalignment of orientation between the current and goal positions. Finally, a home vector between these two positions is calculated using the SIFT matches as a correspondence field [42]. Another original paper using SIFT matching presents a method for search and localization of objects with a mobile robot. The recognition mechanism of the method is based on receptive field cooccurrence histograms and SIFTS matching [43]. Guzel and Bicker have recently proposed a new SIFT-based algorithm which is inspired from a visual homing strategy and adapted to a behavioural architecture to overcome goal-based navigation problems [44].

Another popular robust local feature detector is speeded up robust features (SURF) which is first presented by Bay et al. in 2003. It is mainly inspired by the SIFT descriptor. However, the standard version of SURF is several times faster than SIFT [45]. Pictographs are widely used in a general environment. Accordingly, a novel navigation system has been recently developed for indoor service using SURF algorithm to detect pictographs [46]. A similar study was presented by Hirose et al. in 2011, detecting pictographs with SURF for robot navigation [47].

In an alternative study, SURF algorithm is used to improve the performance of appearance-based localization methods. Experiments were conducted by using several omnidirectional images [48]. Other corresponding studies can be found in references [49, 50].

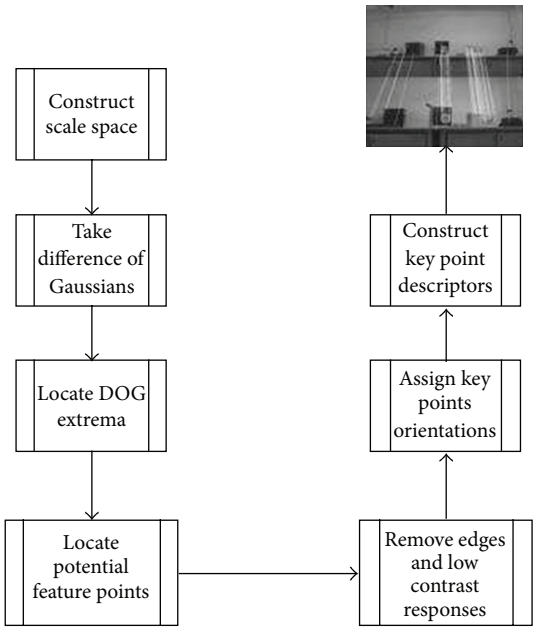


Figure 3: Overview of the SIFT algorithm [39, 40].

A comprehensive research has been recently presented by Martinez-Garcia and Torres [51]. They developed an indoor/outdoor visual odometry scheme for a wheeled mobile robot based on a single passive vision sensor. The authors propose an algorithm which is able to track multiple naturally selected landmarks featured by regional descriptors. The algorithm is in essence an improved version of the maximally stable extremal regions descriptor (MSER). Furthermore, this study was improved by evaluatıng other popular invariant descriptors such as SIFT, Quick SHIFT and feature points extraction algorithm Harris corner detection (HCD) [52], and the fast corner detection (FCD) [53]. Experimental results performed in both indoor and outdoor environments with a mobile robot prove the robustness and consistency of the proposed algorithm. One of the most important criteria in feature tracking-based navigation systems is the computational time consumed by the algorithm. Accordingly, researches have been focused on performance improvement in corresponding algorithms. The most efficient way of performance increment is to parallelize the code or to execute the application on multiprocessor architectures. Various techniques have been proposed to increase parallelism of streaming applications. For instance, Wang et al. have recently presented a comprehensive method in order to reduce schedule length and energy consumption on multiprocessor system-on-chips (MPSoCs). They proposed an ILP design to generate an optimum objective task schedule [54]. Alternatively OpenMP architecture is a robust way of accelerating existing invariant descriptors [44, 55, 56]. The architecture generates multiple threads for parallel programming in a public, shared memory environment.

Visual servoing is another important concept which can be included in this category and is defined as the capability to employ visual information to control the pose of the robot's

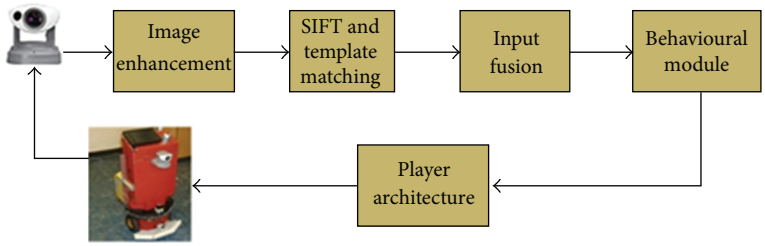


Figure 4: Control architecture of the obstacle avoidance algorithm [64].

end-effectors relative to a target object or a set of target features. The task can also be defined for mobile robots, where it becomes the control of the vehicle's pose with respect to specific landmarks. Thus, Szenher defined the feature-based visual homing algorithms as a type of image-based visual servoing [31]. There are two main approaches for visual servoing systems, namely: position-based visual servoing (PBVS) and image-based visual servoing (IBVS) [57]. PBVS algorithms solve the trajectory problem in workspace. However, in IBVS, the control commands are deduced directly from image features. Kim and Oh have proposed an intelligent mobile robot navigation architecture comprising both of these servoing methods to guide a mobile robot [58]. The IBVS module estimates the motion planning directly from the image space so as to keep the target object always in the field of view. As well as this, the PBVS module is employed to conduct an image-to-workspace transform to plan an optimal pose trajectory directly in the Cartesian space. Another study focuses on docking an AUV to an underwater station with a single standard camera and visual servoing strategy [59].

3.5. Obstacle Avoidance Systems Using Qualitative Information. Obstacle avoidance techniques classified essentially entail extraction of qualitative image characteristics and their interpretation [6]. They are basically defined as sensor-based obstacle avoidance systems which process every item of online sensor data to estimate free and occupied space. These methods are considered useful in avoiding having to compute accurate numerical data such as distance and position coordinates. Lorigo et al. proposed a low resolution vision-based obstacle avoidance architecture consisting of three-dependent vision modules for obstacle detection. These modules were associated with edges, red, green, and blue (RGB) colours, and hue, saturation, and value (HSV) information. The data from these modules was analyzed by a fourth module so as to simultaneously generate motion commands [60]. Ulrich and Nourbakhsh proposed a similar vision-based obstacle avoidance strategy based on monocular vision. The strategy essentially involves assigning each pixel as either obstacle or ground according to its colour appearance [61].

Saitoh et al. integrated this obstacle avoidance technique into a centre followed-based mobile robot navigation architecture. The system does not need prior knowledge of the environment and employs a low cost monocular vision camera as the only sensor needed to navigate the robot safely [62]. The robot has a basic navigation strategy so that it moves towards the centre of the corridor until it encounters an unexpected obstacle. When any obstacle is detected, the robot attempts to avoid it or stops depending on the size of the obstacle. If the robot manages to pass the obstacle successfully, it then localizes itself toward the centre.

ROBOCUP competition has become quite popular and has attracted the attention of many researchers in recent years. The detection of an opponent robot and the ball are two challenging tasks which must be solved efficiently. Fasola and Veloso proposed using image colour segmentation techniques for object detection and grayscale image processing to detect the opponent robots [63]. Guzel and Bicker proposed a new method inspired from feature matching principal and combined Sift-based feature detection and conventional appearance-based method using a fuzzy inference system [64]. The control architecture of the proposed navigation architecture is illustrated in Figure 4.

NASA's twin Mars Exploration Rovers (MERs), Spirit and Opportunity uses a local path planner (GESTALT) equipped with stereo cameras to evaluate terrain safety avoid obstacles [65]. Despite that GESTALT performed well when guiding the rovers around narrow and isolated hazards, it failed when clusters of closely spaced, nontraversable rocks form extended obstacles. To address this problem, a new technology was proposed at the Jet Propulsion Laboratory. According to which Field D* global path planner has been integrated into rover's software which allows simultaneous local and global planning during navigation. This resulted in better obstacle avoidance and navigation around hazards [66].

A comprehensive review paper about this issue can be found in [67]. The paper introduces major milestones in the development of computer vision for autonomous vehicles over the last decades. Besides, it discusses the design and performance of computer vision algorithms used on Mars in the NASA/JPL Mars Exploration Rover (MER) mission. The paper reveals that the MER mission has been achieved far beyond its primary objectives. The three computer vision algorithms developed for the mission, namely: stereo vision, visual odometry for the rover, and feature tracking for the lander. Despite the limitations of processors deployed, they performed consistently and made significant contributions to the project.

This section details the algorithms and method using mapless strategies for vision-based robot navigation. To summarize this approach, it can be stated that mapless strategies

Table 1: Summary of milestones in visual navigation.

Authors	Category	Strategy	Type of vehicle	Type of sensor
[8]	Map based	Absolute	Ground	Single camera
[9]	Map based	Incremental	Ground	Single camera
[10, 11]	Map based	Incremental	Ground	Stereo vision
[12]	Map based	Landmark	Ground	Stereo vision
[16, 17]	Map building	Slam	Ground	Stereo vision
[19, 20]	Mapless	Optical flow	Ground	Stereo camera
[21, 22]	Mapless	Optical flow	Ground	Single camera
[25, 26]	Mapless	Optical flow	UAV	Single camera
[27, 28]	Mapless	Appearance based	Ground	Single camera
[29, 31]	Mapless	Visual homing	Ground	Omnidirectional
[33, 34]	Mapless	Object recognition	Ground	Single camera
[35, 38, 41, 43, 47, 51]	Mapless	Feature tracking	Ground	Single camera
[48]	Mapless	Feature tracking	Ground	Omnidirectional
[61, 63, 65]	Mapless	Qualitative avoidance	Ground	Single camera
[59]	Mapless	Visual servoing	AUV	Single camera

do not require any explicit representation of the working environment where navigation is to occur.

Principally, these strategies involve recognizing objects found in the environment or tracking those objects using visual clues or observations. Alternatively, motion estimation techniques are widely used in this category. Section two, on the other hand, addresses the strategies using or constructing map for navigation. First, map-based techniques have been introduced which basically employs geometric models or topological maps of the environment. Despite providing fast, robust, and consistent solutions to many problems, they are highly dependent on static maps of working environments which limits the operational capability of the algorithms in this group. Finally, map-building-based strategies have been discussed in which sensors are used to construct geometric models or topological model of the environment. These maps are then used to navigate the robot through corresponding environments. Map-building-based strategies can allow autonomous vehicles to navigate through dynamic environments. However, algorithms used in this group cost a lot of time and effort to obtain a robust model of the environment. Besides, vision-based modelling is challenging, comparing with other sensor systems.

Overall, despite challenges in designing and implementation, mapless strategies developed so far resemble human behaviour more than other approaches and have become applicable to any indoor environment consisting of corridors and doorways.

4. Conclusions

Vision is potentially the most powerful sensing capability in providing reliable and safe navigation. Vision-based navigation techniques are fundamentally classified based on localization techniques. This paper has surveyed each group individually, and as well as vision-based obstacle avoidance techniques are placed under a new section. This paper, inspired from two detailed literature study [5, 6], is essentially the selection of the most outstanding contributions surveyed in the area of vision-based robot navigation. One of the most important differences of this study is that this paper mainly focuses on mapless navigation strategies which are not highlighted and detailed in previously published surveys. Table 1 shows an overview and the most outstanding publications referenced in this survey.

References

[1] P. Nattharith, *Mobile robot navigation using a behavioural control strategy [Ph.D. thesis]*, Newcastle University, Newcastle, UK, 2010.

[2] N. J. Nilsson, *Shakey the Robot*, Artifical Intelligence Center, Computer Science and Techonlogy Division, Standford Research Institute, Menlo Park, Calif, USA, 1984.

[3] T. Hellström, *Biological Foundations of Robot Behavior*, pp. 4–6, Department of Computing Science Umeå University, Umeå, Sweden, 2011.

[4] H. P. Moravec, "The Stanford cart and the CMU rover," *Proceedings of the IEEE*, vol. 71, no. 7, pp. 872–884, 1983.

[5] G. N. DeSouza and A. C. Kak, "Vision for mobile robot navigation: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 2, pp. 237–267, 2002.

[6] F. Bonin-Font, A. Ortiz, and G. Oliver, "Visual navigation for mobile robots: a survey," *Journal of Intelligent and Robotic Systems*, vol. 53, no. 3, pp. 263–296, 2008.

[7] J. J. Guerrero, R. Martinez-Cantin, and C. Sagüés, "Visual mapless navigation based on homographies," *Journal of Robotic Systems*, vol. 22, no. 10, pp. 569–581, 2005.

[8] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte Carlo localization for mobile robots," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '99)*, vol. 2, pp. 1322–1328, May 1999.

[9] A. Kosaka and A. Kak, "Fast vision-guided mobile robot navigation using model-based reasoning and prediction of uncertainties," in *Proceedings of the lEEE/RSJ International*

Conference on Intelligent Robots and Systems (IROS '92), July 1992.

[10] L. Matthies and S. A. Shafer, "Error modeling in stereo navigation," *IEEE Journal of Robotics and Automation*, vol. 3, no. 3, pp. 239–248, 1987.

[11] H. I. Christensen, N. O. Kirkeby, S. Kristensen, L. Knudsen, and E. Granum, "Model-driven vision for in-door navigation," *Robotics and Autonomous Systems*, vol. 12, no. 3-4, pp. 199–207, 1994.

[12] M. Hashima, F. Hasegawa, S. Kanda, T. Maruyama, and T. Uchiyama, "Localization and obstacle detection for a robot for carrying food trays," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robot and Systems (IROS '97)*, pp. 345–351, September 1997.

[13] M. R. Kabuka and A. E. Arenas, "Position erification of a mobile robot using standard pattern," *IEEE Journal of Robotics and Automation*, vol. 3, no. 6, pp. 505–516, 1987.

[14] N. Sawasaki, T. Morita, and T. Uchiyama, "Design and implementation of high-speed visual tracking system for real-time motion analysis," in *Proceedings of the IAPR International Conference Pattern Recognition*, vol. 3, pp. 478–483, 1996.

[15] C. E. Thorpe, *An Analysis of Interest Operators for FIDO*, Departmant of Computer Science and The Robotics Institute, Carnegie-Mellon University, Pittsburgh, Pa, USA, 1984.

[16] S. Se, D. Lowe, and J. Little, "Vision-based mobile robot localization and mapping using scale-invariant features," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '01)*, May 2001.

[17] K. Kidono, J. Miura, and Y. Shirai, "Autonomous visual navigation of a mobile robot using a human-guided experience," *Robotics and Autonomous Systems*, vol. 40, no. 2-3, pp. 124–132, 2002.

[18] R. Sim and J. J. Little, "Autonomous vision-based robotic exploration and mapping using hybrid maps and particle filters," *Image and Vision Computing*, vol. 27, no. 1-2, pp. 167–177, 2009.

[19] A. Bernardino and J. Santos-Victor, "Visual behaviours for binocular tracking," *Robotics and Autonomous Systems*, vol. 25, no. 3-4, pp. 137–146, 1998.

[20] A. P. Duchon, W. H. Warren, and L. P. Kaelbling, "Ecological robotics," *Adaptive Behavior*, vol. 6, pp. 1–30, 1994.

[21] S. Temizer, *Optical flow based local navigation [M.S. thesis]*, Massachusetts Institute of Technology, Cambridge, Mass, USA, 2001.

[22] C. McCarthy and N. Barnes, "Performance of optical flow techniques for indoor navigation with a mobile robot," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '04)*, vol. 5, pp. 5093–5098, usa, April 2004.

[23] K. Souhila and A. Karim, "Optical flow based robot obstacle avoidance," *International Journal of Advanced Robotic Systems*, vol. 4, no. 1, pp. 13–16, 2007.

[24] M. S. Guzel and R. Bicker, "Optical flow based system design for mobile robots," in *Proceedings of the IEEE International Conference on Robotics, Automation and Mechatronics (RAM '10)*, pp. 545–550, Singapore, June 2010.

[25] B. Herissé, T. Hamel, R. Mahony, and F. Russotto, "Landing a VTOL unmanned aerial vehicle on a moving platform using optical flow," *IEEE Transactions on Robotics*, vol. 28, no. 1, pp. 77–89, 2012.

[26] S. Temizer and L. P. Kaelbling, *Optical Flow Based Local Navigation*, MIT, Cambridge, Mass, USA, 2003.

[27] P. Gaussier, C. Joulain, S. Zrehen, J. P. Banquet, and A. Revel, "Visual navigation in an open environment without map," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robot and Systems (IROS '97)*, pp. 545–550, September 1998.

[28] Y. Matsumoto, M. Inaba, and H. Inoue, "Visual navigation using view-sequenced route representation," in *Proceedings of the 13th IEEE International Conference on Robotics and Automation (ICRA '96)*, pp. 83–88, April 1996.

[29] C. Zhou, Y. Wei, and T. Tan, "Mobile robot self-localization based on global visual appearance features," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '03)*, pp. 1271–1276, September 2003.

[30] C. Zhichao and S. T. Birchfield, "Qualitative vision-based mobile robot navigation," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '06)*, pp. 2686–2692, May 2006.

[31] D. M. Szenher, *Visual homing in dynamic indoor enviorenments [Ph.D. thesis]*, University of Edinburgh, Edinburgh, UK, 2008.

[32] P. Arena, S. de Fiore, L. Fortuna, L. Nicolosi, L. Patané, and G. Vagliasindi, "Visual homing: experimental results on an autonomous robot," in *Proceedings of the 18th European Conference on Circuit Theory and Design 2007 (ECCTD '07)*, pp. 304–307, August 2007.

[33] D. Kim and R. Nevatia, "Recognition and localization of generic objects for indoor navigation using functionality," *Image and Vision Computing*, vol. 16, no. 11, pp. 729–743, 1998.

[34] D. Kim and R. Nevatia, "Symbolic navigation with a generic map," *Autonomous Robots*, vol. 6, no. 1, pp. 69–88, 1999.

[35] R. C. Harrell, D. C. Slaughter, and P. D. Adsit, "A fruit-tracking system for robotic harvesting," *Machine Vision and Applications*, vol. 2, no. 2, pp. 69–80, 1989.

[36] P. E. Trahanias, S. Velissaris, and T. Garavelos, "Visual landmark extraction and recognition for autonomous robot navigation," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robot and Systems (IROS '97)*, pp. 1036–1042, September 1998.

[37] A. A. Argyros and F. Bergholm, "Combining central and peripheral vision for reactive robot navigation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '99)*, pp. 646–651, June 1999.

[38] H. Li and S. X. Yang, "A behavior-based mobile robot with a visual landmark-recognition system," *IEEE/ASME Transactions on Mechatronics*, vol. 8, no. 3, pp. 390–400, 2003.

[39] M. C. Deans, *Bearings-only localization and mapping [Ph.D. thesis]*, Mellon University, Pittsburgh, Pa, USA, 2005.

[40] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV '99)*, pp. 1150–1157, IEEE, Kerkyra, Greece, September 1999.

[41] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[42] J. S. Pons, W. Hübner, H. Dahmen, and H. A. Mallot, "Vision-based robot homing in dynamic environments," in *Proceedings of the 13th IASTED International Conference on Robotics and Applications*, pp. 293–298, ACTA Press, Wuerzburg, Germany, August 2007.

[43] K. Sjöö, G. D. Lopez, C. Paul, P. Jensfelt, and D. Kragic, "Object search and localization for an indoor mobile robot," *Journal of Computing and Information Technology*, vol. 17, no. 1, pp. 1–12, 2009.

[44] M. S. Guzel and R. Bicker, "A behaviour-based architecture for mapless navigation using vision," *International Journal of Advanced Robotic Systems*, vol. 9, pp. 1–13, 2012.

[45] H. Bay, A. Ess, T. Tuytelaars, and L. van Gool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2003.

[46] D. Chugo, K. Hirose, K. Nakashima, S. Yokota, H. Kobayashi, and H. Hashimoto, "Camera-based navigation for service robots using pictographs on the crossing point," in *Proceedings of the 38th Annual Conference on IEEE Industrial Electronics Society*, Monteral, Canada, October 2012.

[47] K. Hirose, D. Chugo, S. Yokota, and K. Takase, "Service robots navigation using pictographs detection for indoor environment," in *Proceedings of the 37th Annual Conference of the IEEE Industrial Electronics Society (IECON '11)*, pp. 2170–2175, November 2011.

[48] T. Tasaki, S. Tokura, T. Sonoura, F. Ozaki, and N. Matsuhira, "Mobile robot self-localization based on tracked scale and rotation invariant feature points by using an omnidirectional camera," in *Proceedings of the lEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '10)*, October 2010.

[49] K. Wetzelsberger, T. Pfannschmidt, and T. Ihme, "Real-time image processing for motion planning based on realistic sensor data," in *Proceedings of the 8th IEEE International Workshop on Robotic and Sensors Environments (ROSE '10)*, pp. 158–163, October 2010.

[50] Y. Hagiwara, T. Shoji, and H. Imamura, "Position and rotation estimation for mobile robots in outside of recording path using ego-motion," *IEEJ Transactions on Electronics, Information and Systems*, vol. 133, no. 2, pp. 356–364, 2013.

[51] E. Martinez-Garcia and D. Torres, "Robot visual odometry by tracking multiple natural landmarks with invariant descriptors," ROSSUM, 2011.

[52] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of the 4th Alvey Vision Conference*, pp. 147–151, August 1988.

[53] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," in *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05)*, pp. 1508–1511, October 2005.

[54] Y. Wang, D. Liu, Z. Qin, and Z. Shao, "Optimally removing intercore communication overhead for streaming applications on MPSoCs'computers," *IEEE Transactions*, vol. 62, no. 2, pp. 336–350, 2013.

[55] S. Warn, W. Emeneker, J. Cothren, and A. Apon, "Accelerating SIFT on parallel architectures," in *Proceedings of the IEEE International Conference on Cluster Computing and Workshops (CLUSTER '09)*, pp. 1, 4, September 2009.

[56] S. Jia, X. Yin, and X. Li, "Mobile robot parallel PF-SLAM based on OpenMP," in *Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO '12)*, pp. 508–513, 2012.

[57] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 5, pp. 651–670, 1996.

[58] S. Kim and S. Oh, "Hybrid position and image based visual servoing for mobile robots," *Journal of Intelligent and Fuzzy Systems*, vol. 18, no. 1, pp. 73–82, 2007.

[59] P. Lee, B. Jeon, and S. Kim, "Visual servoing for underwater docking of an autonomous underwater vehicle with one camera," in *Proceedings of the OCEANS 2003*, vol. 2, pp. 677–682, San Diego, Calif, USA, September 2003.

[60] L. M. Lorigo, R. A. Brooks, and W. E. L. Crimson, "Visually-guided obstacle avoidance in unstructured environments," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robot and Systems (IROS '97)*, vol. 1, pp. 373–379, September 1997.

[61] I. Ulrich and I. Nourbakhsh, "Appearance-based obstacle detection with monocular color vision," in *Proceedings of the AAAI NationalConference on Artificial Intelligence*, September 2000.

[62] T. Saitoh, N. Tada, and R. Konishi, "Indoor mobile robot navigation by central following based on monocular vision," *IEEJ Transactions on Electronics, Information and Systems*, vol. 129, no. 8, pp. 1576–1584, 2009.

[63] J. Fasola and M. Veloso, "Real-time object detection using segmented and grayscale images," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '06)*, pp. 4088–4093, May 2006.

[64] M. S. Guzel and R. Bicker, "Vision based obstacle avoidance techniques," in *Recent Advances in Mobile Robotics*, pp. 83–108, InTech, New York, NY, USA, 2012.

[65] J. Carsten, A. Rankin, D. Ferguson, and A. Stentz, "Global path planning on board the mars exploration rovers," in *Proceedings of the 2007 IEEE Aerospace Conference*, March 2007.

[66] J. Carsten, A. Rankin, D. Ferguson, and A. Stentz, "Global planning on the mars exploration rovers: software integration and surface testing," *Journal of Field Robotics*, vol. 26, no. 4, pp. 337–357, 2009.

[67] L. Matthias, M. Maimone, and A. Johnson, "Computer vision on Mars," *International Journal of Computer Vision*, vol. 75, no. 1, pp. 67–92, 2007.