Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.777635

RECENT ADVANCES IN COMPUTER VISION APPLICATIONS OBJECT RECOGNITION AND DEEP LEARNING

Jasmin Praful Bharadiya*
*Affiliation: Department of Information and Technology, University of the Cumberlands, Williamsburg, KY USA, Email: jasminbharadiya92@gmail.com. ORCID: https://orcid.org/0009-0002-4264-6005

***Corresponding Author:** Jasmin Praful Bharadiya
*Affiliation: Department of Information and Technology, University of the Cumberlands, Williamsburg, KY USA, Email: jasminbharadiya92@gmail.com. ORCID: https://orcid.org/0009-0002-4264-6005

Abstract:
Moving object detection has achieved a noticeable attention in many computer vision applications. The research community have contributed lot of works for dealing with major challenges of moving object detection in real-world scenarios. The paper presents a comprehensive review on different moving object detection techniques classified into four categories: Background Modeling Based techniques; Frame Difference Based techniques; Optical Flow Based techniques and Deep Learning Based techniques. Moreover, detailed descriptions of various methods in each of this category are also provided.
Keywords: Computer Vision; Moving Object Detection Techniques; Comprehensive Survey.

Introduction:

During the last few decades, automated video analysis has become a potential research area in computer vision due to its numerous applications to video based intelligent systems. There are three fundamental stages for analysis of video sequences: detection of salient moving objects, tracking of these salient objects on frame basis, and analysis of object tracks to predict the activity or behaviour of this objects. Furthermore, surveillance system have significant role in the defence against criminality and terrorist threats in both public and private sectors. It rely on the ability to detect moving objects in outdoor and indoor scenes which is considered as an efficient step for information extraction in computer vision applications. The term 'object' usually refers to its generalized form, including pedestrians and man-made objects (e.g. vehicles, ships, buildings, etc.) that have sharp boundaries and are independent of background environment.

However moving object/foreground object detection has been a challenging task because of the following reasons: Multiple moving objects may present in the scene; small and poorly textured moving objects; poor and rapid change in illumination conditions and shadows and multiple occlusions exist. During the last decades, considerable efforts have been made to develop various methods for the detection of different types of moving objects depicting vehicles and pedestrians in indoor or outdoor scenes objects. While enormous methods exist, a deep review and experimental analysis of the literature concerning generic object detection

is still lacking. Depending upon this phenomenon, the paper presents a critical review on existing models for detection of moving objects in indoor or outdoor conditions. In the present survey, we categorized the object detection techniques depending on the approaches they used to detect the salient moving objects. This survey will be especially beneficial for the researchers to have better understanding of this research field and select the most suitable algorithm for their particular needs.
In the realm of computer vision, one of the most captivating and dynamically evolving subfields is that of moving object detection. The capacity to perceive and track objects in motion is pivotal, finding applications across a spectrum of domains, including surveillance systems, robotics, autonomous vehicles, and augmented reality. The complexities and challenges associated with moving object detection in real-world scenarios have fueled a prolific body of research, leading to the development of an array of innovative techniques and methodologies.

This article embarks on a comprehensive exploration of the multifaceted landscape of moving object detection. It is characterized by a nuanced understanding of the field's intrinsic challenges and the ingenious solutions that researchers have devised over time. Specifically, this review categorizes these solutions into four distinct classes, each representative of different paradigms: Background Modeling Based techniques, Frame Difference Based techniques, Optical Flow Based techniques, and Deep Learning Based techniques. These categories serve as organizational pillars, allowing for a structured examination of the state-of-the-art in moving object detection.

The significance of this survey lies not only in its capacity to provide a bird's-eye view of the field's current standing but also in its potential to guide and inform both seasoned researchers and newcomers. With the evolving diversity of moving object detection approaches, selecting the most apt algorithm for a particular application can be a formidable task. This review seeks to alleviate this challenge by offering a comprehensive panorama, enabling stakeholders to navigate through the various methodologies with greater clarity and discernment. As the evolution of computer vision continues to advance, propelled by emerging technologies and ever-expanding applications, understanding the underpinnings of moving object detection is paramount. This review, with its categorization and exploration of techniques, as well as its insights into the nuances of each category, contributes to the broader discourse within computer vision and fosters a deeper appreciation of the profound impact that moving object detection has on the world of visual perception and intelligent systems.

In the rapidly evolving field of computer vision, the pursuit of robust moving object detection has taken center stage, driven by its pivotal role in a myriad of real-world applications. Whether it's enhancing surveillance systems, enabling autonomous vehicles to navigate complex environments, or powering interactive augmented reality experiences, the ability to accurately identify and track moving objects is foundational.

The intricacies and challenges associated with moving object detection in dynamic, real-world scenarios have ignited a flurry of research activity, resulting in a diverse array of innovative techniques and methodologies. This article embarks on a comprehensive exploration of this vibrant landscape, shedding light on both the inherent complexities and ingenious solutions that researchers have crafted over time. To provide clarity and structure to this review, we categorize these solutions into four distinct paradigms, each representing a unique approach to moving object detection: Background Modeling Based techniques, Frame Difference Based techniques, Optical Flow Based techniques, and Deep Learning Based techniques. Let's illustrate these categories with examples**:**

Background Modeling Based Techniques:

Example: The Gaussian Mixture Model (GMM) is a classic background modeling method. It characterizes the background of a video scene as a mixture of Gaussian distributions. Any deviation from this learned background model is considered a foreground object. GMM has been widely used in early moving object detection systems.

Frame Difference Based Techniques:

Example: Frame differencing involves subtracting consecutive frames in a video sequence to highlight areas of change. If the difference exceeds a predefined threshold, those areas are considered as moving objects. Frame differencing is computationally efficient and has been used in applications like traffic monitoring to detect vehicle movement.

Optical Flow Based Techniques:

Example: Optical flow methods estimate the motion of objects by analyzing pixel-level motion patterns between frames. Lucas-Kanade and Horn-Schunck are two classic optical flow algorithms. These techniques are used in video tracking systems, where understanding object motion is crucial.

Deep Learning Based Techniques:

Example: Convolutional Neural Networks (CNNs) have revolutionized moving object detection. Models like YOLO (You Only Look Once) and Faster R-CNN use deep learning to simultaneously detect and track objects with high accuracy. YOLO, for instance, can detect objects in real-time and is used in applications like object recognition for autonomous vehicles.

The significance of this review lies not only in its capacity to provide a comprehensive overview of the field's current state but also in its potential to guide researchers, practitioners, and enthusiasts. With the ever-expanding diversity of moving object detection approaches, selecting the most suitable algorithm for a specific application can be a daunting task. This review aims to provide the necessary insights and context to navigate through these methodologies effectively, fostering a deeper understanding of the profound impact of moving object detection on the broader landscape of computer vision.

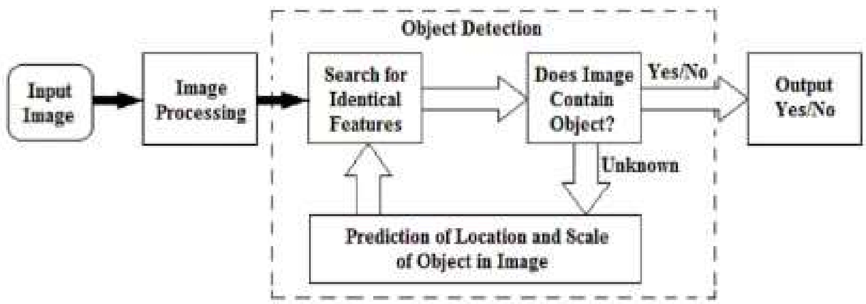

Fig: Object detection

General methodology

The OD system basically comprises of two main phases namely: the learning phase and the testing phase which are shown in Figure 2 that shows the normal working of the OD system. Learning phase is mainly meant for the classifier so that it recognises the objects present in the image that is given as input to the system. Learning phase can be further classified as learning through training and learning through validation. Learning through training comprises mainly of the learning block where a proper learning scheme is defined, it can be part-based or patch-based, etc. The object template block then makes use of the learning’s that were done previously to represent the objects with various representations like histogram representation, random forest representation, etc. Whereas on the other hand, learning through validation block does not require any sort of training as they are validated beforehand. Hence after preprocessing the image, directly template matching is done which produces the features of an object in the image. The main purpose of the testing phase is to decide whether an object is present in the image that is given to the system as input and if yes then to which object class does it belongs to. Here the image is searched for an object by various searching techniques like the sliding window technique, and according to the output of the searching mechanism, a decision is made on the object class.

Classification of OD mechanisms This section classifies the various OD mechanisms based on search, feature classification, template creation and based on matching. We have classified the OD types as sliding window-based, contour-based, graph-based, fuzzy-based, contextbased and some other types. Here we will review the work carried out by various authors in the field of OD. 3.1 Sliding window-based OD Sliding window OD has received remarkable attention as it is considered as a very basic method of detecting objects in an image or video. The sliding window technique basically works by searching through the whole image or scene in order to find the object that is of interest. This is the reason why it failed to meet the criteria of real time applications due to higher execution times and inaccurate localisation. Localisation accuracy is important especially while the OD process is to be followed by object recognition. Bergboer et al. (2007) have studied the various learning methods which are use for realising context-based OD in paintings, namely the gradient method and the context detection method. The gradient method is used to transform a spatial context into a gradient towards an object, whereas the context detection method makes use of the sliding window approach to search the image regions that are likely to contain the object of interest. Basically the gradient method works totally on assumptions which may lead to higher timing constraints when there is only single object that is to be detected in an image. On the other hand, the context detection works based on sliding window which again introduces the timing constraint as it searches each window for the presence of an object. Clearly the issue of inaccurate localisation has been a concern while using the sliding window-based OD technique; Segvic et al. (2011) have explained how localisation accuracy could be achieved by removing the need for spatial clustering of the nearby detection responses. This leads to three main goals namely high recall, high precision and accurate localisation. Spatial clustering could be used to suppress the number of false positives but at the price of localisation uncertainty. Sliding window technique initially fixes the size of the window in which it will be searching for the object, but in order to increase the rate at which the detections must happen, Comaschi et al. (2013) have proposed a sliding window approach that decides on the step size of the window at run time, which helps to apply this technique of sliding window to real time applications. They have also demonstrated that how this technique improves the performance of Viola Jones OD, and also claimed to have achieved a speedup of 2.03x in frames per second without compromising the accuracy factor. The main issue being the space utilised. Divvala (2012) has studied the two factors that influence the performance of sliding window technique for OD, namely context and subcategories. The use of the first factor that is context shows how the performance of the sliding window approach could be improved. As the siding window approach searches the total image for the presence of an object in it, while the use of context can be made in order to know whether a particular object is present in that region or not. The subcategories factor is where the information within the sliding window is used to split the training data into smaller groups which have reduced appearance diversity which leads to simpler classification. Here he has discussed only about the two factors, there could be many other factors too such as contour which could affect the performance to a great deal. Shave presented a technique which is used to reduce the number of miss detections while increasing the grid spacing while the sliding window approach is used for OD. They have achieved it by using a patch to predict the bounding box of an object within the search area, and in order to improve the speed of estimating the bounding boxes the authors make use of the decision tree with simple binary test at each node. Although they claim that their proposed system works on a wide variety of images, still we feel that an occluded image could remain a challenge.

Future Challenging's:
The future challenges in object detection within computer vision:

Real-Time Object Detection at Scale:
Challenge: Real-time object detection, especially for high-resolution videos, demands substantial computational resources. Achieving real-time performance without compromising accuracy remains a formidable challenge.

Solution: Developing efficient network architectures, hardware acceleration (e.g., GPUs and TPUs), and model quantization techniques are avenues to address this challenge.

Handling Extremely Large Datasets:
Challenge: The growth of deep learning and the availability of massive datasets pose challenges in terms of memory, storage, and computational requirements for training object detection models.

Solution: Distributed training, data augmentation techniques, and strategies for selecting representative subsets of data are vital for scaling up object detection.

Robustness to Adverse Conditions:
Challenge: Objects in the real world often encounter challenging conditions such as poor lighting, adverse weather, occlusions, and complex backgrounds. Object detectors must remain robust under these circumstances.

Solution: Data augmentation with synthetic adverse conditions, sensor fusion (e.g., combining RGB and infrared data), and advanced pre-processing techniques can enhance robustness.

Generalization Across Object Categories:
Challenge: Object detectors should generalize well to a wide variety of object categories, including rare or previously unseen ones, with limited labeled data.

Solution: Few-shot learning, transfer learning, and meta-learning techniques aim to improve the model's ability to adapt to new object categories with minimal annotated samples.

Reducing Annotation Effort:
Challenge: Manually annotating objects in images is expensive and time-consuming. Reducing the annotation effort while maintaining detection accuracy is a pressing challenge.

Solution: Weakly supervised learning, active learning, and self-supervised learning are approaches to reduce the need for extensive manual annotation.
Interpretability and Explainability:

Challenge: In safety-critical applications, understanding why an object detector made a particular decision is crucial. Black-box models lack interpretability.

Solution: Developing models that provide interpretable outputs, such as attention mechanisms and visual explanations, helps users understand model decisions.

Privacy and Ethical Concerns:
Challenge: Object detection in public spaces raises ethical concerns about privacy, surveillance, and potential misuse.

Solution: Striking a balance between the benefits and ethical considerations through legislation, privacy-preserving techniques (e.g., federated learning), and transparent deployment practices is essential.

Handling 3D Objects:
Challenge: In applications like robotics and augmented reality, detecting and recognizing objects in three-dimensional space poses new challenges.

Solution: Combining 2D and 3D object detection techniques, incorporating depth information (e.g., from LiDAR or depth cameras), and advancing 3D object detection networks are crucial.

Adversarial Attacks:
Challenge: Adversarial attacks can trick object detectors into making incorrect predictions. Robustness against such attacks is crucial.

Solution: Adversarial training, model robustness evaluations, and incorporating adversarial detection mechanisms help mitigate these threats.

Energy-Efficient Object Detection:
Challenge: In resource-constrained environments, energy-efficient object detection is vital. Optimizing models for low-power consumption without sacrificing accuracy is challenging.

Solution: Model pruning, quantization, and hardware optimizations tailored for energy-efficient inference are key solutions.

Cross-Modal Object Detection:
Challenge: Integrating data from different sensors (e.g., cameras and LiDAR) for improved object detection in applications like autonomous vehicles requires robust cross-modal techniques.

Solution: Multi-modal fusion strategies, sensor calibration, and cross-modal object detection models enable effective integration of sensor data.

Addressing these detailed challenges necessitates interdisciplinary collaboration among computer vision researchers, hardware experts, ethicists, and policymakers. Moreover, ongoing discussions around responsible AI deployment are essential to ensure that object detection technologies align with societal values and ethical principles.

Recommendation:

To advance the field of object detection, a multifaceted approach is crucial. First and foremost, substantial investment in research and development should be made, fostering innovation through collaboration between academia, industry, and government entities. Open access to diverse datasets and standardized evaluation metrics is essential to ensure fair comparisons and benchmarking. Efficiency is a paramount concern, demanding the design and optimization of model architectures for real-time detection across various hardware platforms, alongside exploring quantization and neural architecture search for reduced computational demands. Moreover, data efficiency must be improved through techniques like transfer learning and active learning, enabling object detectors to learn from limited annotated data. Robustness, both in terms of handling realworld conditions and defending against adversarial attacks, should be a primary research focus. The development of interpretable models and ethical guidelines is imperative, ensuring transparency and responsible AI practices. Privacy-preserving techniques like federated learning should be explored, particularly in surveillance and data-sharing contexts. Cross-modal integration between computer vision and sensor technology experts is essential for applications like autonomous vehicles. Energy efficiency and public awareness should not be neglected, with hardware optimizations and educational efforts playing critical roles. Establishing regulations and standards for sensitive applications and continuous evaluation and benchmarking platforms are necessary for tracking progress and ensuring responsible deployment. Ultimately, promoting diversity and fostering international collaboration will lead to more inclusive and globally beneficial object detection solutions.

Here are some recommendations and strategies to address the challenges in object detection and to advance the field:

Invest in Research and Development:

Allocate resources for ongoing research to foster innovation in object detection techniques.

Encourage collaboration between academia, industry, and government agencies to support research initiatives.

Open Access to Datasets and Benchmarks:

Promote the creation and sharing of diverse and comprehensive datasets for benchmarking and evaluating object detection algorithms.

Encourage the development of standardized evaluation metrics to facilitate fair comparisons.

Efficient Model Architectures:

Invest in designing and optimizing model architectures for real-time object detection on various hardware platforms.

Explore model quantization and neural architecture search to reduce computational requirements.

Data Efficiency and Transfer Learning:
Develop techniques that enable object detectors to learn from limited annotated data through transfer learning, few-shot learning, and semi-supervised approaches.

Focus on active learning strategies to intelligently select the most informative samples for annotation.

Robustness and Adversarial Defense:
Invest in research on adversarial defense mechanisms to ensure the robustness of object detection systems against attacks.

Continuously evaluate models against diverse and challenging real-world conditions.

Interpretable Models:
Promote research into interpretable deep learning models that provide transparency and insights into model decision-making.

Develop visualization tools and techniques to aid users in understanding model outputs.

Ethical Considerations:
Establish ethical guidelines and codes of conduct for the development and deployment of object detection systems.

Encourage responsible AI practices and consider the societal impact of these technologies.

Privacy-Preserving Techniques:
Explore privacy-preserving methods, such as federated learning and secure multi-party computation, to protect individuals' privacy in surveillance and data-sharing scenarios.

Cross-Modal Integration:
Foster interdisciplinary collaboration between computer vision and sensor technology experts to advance cross-modal object detection for applications like autonomous vehicles.

Energy Efficiency:
Investigate hardware-specific optimizations and model compression techniques to enable energyefficient object detection on edge devices and mobile platforms.

Public Awareness and Education:

Educate the public about the capabilities, limitations, and potential ethical concerns of object detection technologies.

Engage in public discourse and involve stakeholders in shaping regulations and policies.

Regulations and Standards:
Advocate for the development of regulations and standards governing the use of object detection in sensitive applications to ensure ethical and responsible deployment.

Continuous Evaluation and Benchmarking:
Establish platforms for continuous evaluation and benchmarking of object detection models to track progress and identify areas that require improvement.

Diverse Representation:
Promote diversity and inclusivity in object detection research and development to ensure that systems work well for all demographic groups.

International Collaboration:
Encourage collaboration and knowledge-sharing among researchers, organizations, and governments on a global scale to address common challenges and promote responsible AI.

By implementing these recommendations, the field of object detection can continue to advance, delivering more accurate, efficient, and responsible solutions that benefit a wide range of applications while addressing societal concerns.

Conclusion

In conclusion, object detection is a dynamic and vital field within computer vision, with a wide range of applications that impact our daily lives. As we navigate the complexities of real-world scenarios, it becomes evident that numerous challenges and opportunities lie ahead. This comprehensive survey has illuminated the diverse landscape of moving object detection techniques, spanning background modeling, frame differencing, optical flow analysis, and deep learning approaches. It has also underscored the importance of this field in applications like surveillance systems, robotics, autonomous vehicles, and augmented reality.

Looking to the future, it is imperative that we continue to invest in research and development, fostering innovation and collaboration among researchers, industry stakeholders, and policymakers. Open access to datasets and standardized evaluation metrics will enable fair and transparent assessments of object detection algorithms. Efficiency and data efficiency must remain at the forefront of our efforts, with a focus on optimized model architectures, transfer learning, and active learning strategies.

Moreover, addressing the challenges of robustness against adversarial attacks and interpretability for user trust is paramount. We must also consider the ethical implications of object detection technologies, promoting responsible AI practices and privacy-preserving techniques.

As we move forward, cross-modal integration, energy efficiency, public awareness, and international collaboration will be key drivers of progress. Establishing regulations and standards to govern sensitive applications and maintaining continuous evaluation platforms are essential to ensure responsible deployment.

In essence, the future of object detection holds tremendous potential for innovation and positive impact on society. By addressing these challenges and embracing these opportunities, we can continue to advance the field, delivering more accurate, efficient, and ethically sound solutions that empower a wide range of applications and benefit humanity as a whole.

References

[1]. Al-Khalifa, H.: Utilizing QR code and mobile phones for blinds and visually impaired people. In: Computers Helping People with Special Needs, pp. 1065–1069 (2008) Amazon mechanical turk. http://www.mturk.com/

[2]. Ancuti, C., Bekaert, P.: Sift-cch: increasing the sift distinctness by color co-occurrence histograms. In: The Fifth International Symposium on Parallel and Distributed Processing and Applications (ISPA07), pp. 130–135 (2007) Apple Inc. http://www.apple.com/iphone/

[3]. Bakken, T.: An evaluation of the SIFT algorithm for CBIR. Telenor R&I N 30/2007, 2007

[4]. Bauer, J., Sünderhauf, N., Protzel, P.: Comparing several implementations of two recently published feature detectors. In: Proc. of the International Conference on Intelligent and Autonomous Systems, Toulouse, France (2007)

[5]. Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-up robust features (SURF). Comput. Vis. Image Underst. 110, 346–359 (2008)

[6]. Bharadiya , J. P., Tzenios, N. T., & Reddy , M. (2023). Forecasting of Crop Yield using Remote Sensing Data, Agrarian Factors and Machine Learning Approaches. Journal of Engineering Research and Reports, 24(12), 29–44. https://doi.org/10.9734/jerr/2023/v24i12858

[7]. Bharadiya, J. P. (2023). Introduction of the Popular Machine Learning Algorithm. International Journal of Innovative Science and Research Technology, 8(5), 2504-2508. Retrieved from https://bit.ly/3TmGbDi

[8]. Bharadiya, J. P., (2023). Fake News Contrary to Social Media Deep Learning & Machine Learning, Journal of Data Acquisition and Processing, 38 (3): 6950-6958 http://sjcjycl.cn/article/view-2023/03_6950.php

[9]. Blauert, J.: Spatial Hearing: The Psychophysics of Human Sound Localization. MIT Press, Cambridge (1997)

[10]. Canny, J.F.: A computational approach to edge detection. In: Martin, A.F., Oscar, F. (eds.) Readings in Computer Vision: Issues, Problems, Principles, and Paradigms, pp. 184–203. Morgan Kaufmann, San Mateo (1987)

[11]. Chincha, R., Tian, Y.: Finding objects for blind people based on SURF features. In: 2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops, pp. 526–527 (2011)
[12]. Fan, P., Men, A., Chen, M., Yang, B.: Color-SURF: a surf descriptor with local kernel color histograms. Presented at the IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC) 2009, Beijing, China (2009)
[13]. Hasanuzzaman, F.M., Yang, X., Tian, Y.: Robust and effective component-based banknote recognition by SURF features. In: Proc. WOCC, pp. 1–6 (2011)
[14]. Hub, A., Hartter, T., Ertl, T.: Interactive localization and recognition of objects for the blind. In: Northridge Center on Disabilities' 21st
[15]. Annual International Technology and Persons with Disabilities Conference. California State University (2006)
[16]. Hub, A., Hartter, T., Ertl, T.: Interactive tracking of movable objects for the blind on the basis of environment models and perceptionoriented object recognition methods. In: Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility, Portland, Oregon, USA, pp. 111–118 (2006)
[17]. Kawai, Y., Tomita, F.: A support system for visually impaired persons to understand three-dimensional visual information using acoustic interface. In: Proceedings of the 16th International Conference on Pattern Recognition (ICPR'02), vol. 3, pp. 974–977 (2002)
[18]. Kawai, Y., Ueshiba, T., Ishiyama, Y., Sumi, Y., Tomita, F.: Stereo correspondence using segment connectivity. Presented at the Proceedings of the 14th International Conference on Pattern Recognition (1998)
[19]. Kulyukin, V., Kutiyanawala, A.: Eyes-free barcode localization and decoding for visually impaired mobile phone users. In: The 2010 International Conference on Image Processing, Computer Vision, and Pattern Recognition, Las Vegas, NV (2010)
[20]. Kulyukin, V., Kutiyanawala, A.: From ShopTalk to ShopMobile: vision-based barcode scanning with mobile phones for independent blind grocery shopping. In: The 33-rd Annual Conference of the Rehabilitation Engineering and Assistive Technology Society of North America, Las Vegas, Nevada, USA (2010)
[21]. Kutiyanawala, A., Kulyukin, V.: An eyes-free vision-based UPC and MSI barcode localization and decoding algorithm for mobile phones. In: Envision 2010 (2010)
[22]. Lanigan, P.E., Paulos, A.M., Williams, A.W., Rossi, D., Narasimhan, P.: Trinetra: assistive technologies for grocery shopping for the blind. In: 10th IEEE International Symposium on Wearable Computers, Montreux, pp. 147–148 (2006)
[23]. Lanigan, P.E., Paulos, A.M., Williams, A.W., Rossi, D., Narasimhan, P.: Trinetra: assistive technologies for grocery shopping for the blind. In: IEEE-BAIS Symposium on Research in Assistive Technologies, Dayton, OH, pp. 29–36 (2007)
[24]. Leibs, A.: Top 10 iPhone Apps for the Visually Impaired. 22 August 2012
[25]. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 60, 91–110 (2004)
[26]. Lowe, D.G.: Object recognition from local scale-invariant features. In: The Seventh IEEE International Conference on Computer Vision, Kerkyra, pp. 1150–1157 (1999)

[27]. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Trans. Pattern Anal. Mach. Intell. 27, 1615–1630 (2005)
[28]. Mikolajczyk, K., Schmid, C.: Scale and affine invariant interest point detectors. Int. J. Comput. Vis. 60, 63–86 (2004)
[29]. Nicholson, J., Kulyukin, V., Coster, D.: ShopTalk: independent blind shopping through verbal route directions and barcode scans. Open Rehabil. J. 2, 11–23 (2009)
[30]. Pavlidis, T.: An evaluation of the Scale Invariant Feature Transform (SIFT). (25 August 2008)
[31]. Perkins Products: TalkingTag LV multi-purpose voice labels. https://secure2.convio.net/psb/site/Ecommerce/352707661?VIEW_PRODUCT=true&product_id=6862&store_id=1101
[32]. Ramisa, A., Vasudevan, S., Aldavert, D., Toledo, R., Mantaras, R.L.d.: Evaluation of the SIFT object recognition method in mobile robots. In: Proceedings of the 2009 Conference on Artificial Intelligence Research and Development: Proceedings of the 12th International Conference of the Catalan Association for Artificial Intelligence, pp. 9–18 (2009)
[33]. Schauerte, B., Martinez, M., Constantinescu, A., Stiefelhagen, R.: An assistive vision system for the blind that helps find lost things. In: Miesenberger, K., Karshmer, A., Penaz, P., Zagler, W. (eds.) Computers Helping People with Special Needs, vol. 7383, pp. 566–572.
[34]. Springer, Berlin/Heidelberg (2012) Semacode Corporation. http://semacode.com/about/
[35]. Sudol, J., Dialameh, O., Blanchard, C., Dorcey, T.: Looktel—a comprehensive platform for computer-aided visual assistance. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), San Francisco, CA, pp. 73–80 (2010)
[36]. Swain, M., Ballard, D.: Color indexing. Int. J. Comput. Vis. 7, 11–32 (1991)
[37]. Tekin, E., Coughlan, J.M.: A mobile phone application enabling visually impaired users to find and read product barcodes. In: Proceedings of the 12th International Conference on Computers Helping People with Special Needs, Vienna, Austria, pp. 290–295 (2010)
[38]. Winlock, T., Christiansen, E., Belongie, S.: Toward real-time grocery detection for the visually impaired. In: IEEE Computer Society
[39]. Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 49–56 (2010)