Recent Advances in Computer Vision: Applications in Object Recognition and Deep Learning

John Doe
RA, Amazon Web Services, USA

Abstract: The detection of moving objects has become of great interest in numerous computer vision tasks. The research community has extended significant efforts in tackling the problems of detecting moving objects in real-world condition. This paper provides an extensive review of different techniques for moving object detection, which are categorized into four main groups: There are four categories of techniques that have been used in this work: Background Modeling-Based techniques, Frame Difference-Based techniques, Optical Flow-Based techniques, and Deep Learning-Based techniques. Furthermore, the paper provides the specific techniques for each of these categories. **Keywords:** Moving Object Detection Techniques; Computer Vision; Comprehensive Survey.

Introduction:

In the recent years, automated video analysis has become an important research area in computer vision, due to its uses in many video-based intelligent systems. The process of analysing video sequences typically involves three main stages: identifying the existence of important moving objects, tracking these objects in each frame, and interpreting the object motion trajectories to estimate the object's action or behaviours. Furthermore, surveillance systems are of great importance in preventing and mitigating against criminal and terrorist risks in civilian and corporate environments. These systems are based on the capability to recognize moving objects within the indoor and outdoor scenarios, which is crucial to the information acquisition in the computer vision-based applications. In this work, the term 'object' denotes any entity that is bounded, such as persons (including pedestrians) and artifacts (including vehicles, ships, buildings, etc.), regardless of the background.

Due to the various factors like multiple objects in motion in a scene, small objects, low quality object texture, changing light conditions, shadow effects and occlusion, detection of moving or foreground objects has been a difficult task. In the last several decades, much work has been done to address the problem of detecting various kinds of moving objects, such as cars and people, in both indoor and outdoor conditions. Although a large amount of work has been done in this area, there is no systematic survey and experimental evaluation of the literature on generic object detection. This paper seeks to solve this problem by undertaking a critical analysis of the models currently employed for the purpose of identifying moving objects in different situations. The survey categorizes object detection techniques based on the schemes applied for detecting the obvious moving objects. This review should be useful to researchers seeking to learn more about this area and to identify which algorithm may best meet their requirements.

Among all the subfields of computer vision, the detection of moving objects is one of the most exciting and rapidly developing. Object detection and motion analysis are crucial components that are widely used in various fields such as surveillance systems, robotics, autonomous vehicles, etc. The existence of numerous real-world issues with regards to the detection of moving object have led to increased research efforts resulting to more and more new methods and techniques typically involves three main stages: detecting prominent moving objects, tracking these objects frame by frame, and analysing the object tracks to predict their behaviours or activity. Additionally, surveillance systems play a crucial role in protecting against criminal and terrorist threats in both public and private sectors. These systems depend on the ability to detect moving objects in both indoor and outdoor environments, which is an essential step for information extraction in computer vision applications. The term "object" refers to any entity with well-defined boundaries, including both human figures (e.g., pedestrians) and man-made objects (e.g., vehicles, ships, buildings), irrespective of their background environment.

Detecting moving or foreground objects has been a challenging task due to several factors, such as the presence of multiple moving objects in the scene, small or poorly textured objects, variations in lighting conditions, shadows, and occlusions. Over the past few decades, significant efforts have been made to develop methods for detecting different types of moving objects, including vehicles and pedestrians, in both indoor and outdoor environments. Despite the vast number of existing methods, there is still a lack of comprehensive reviews and experimental analysis of literature focused on generic object detection. This paper aims to address this gap by providing a critical review of the current models used for detecting moving objects in various conditions. The survey categorizes object detection techniques based on the approaches used to identify prominent moving objects. This review will be valuable for researchers to gain a deeper understanding of this field and help them choose the most suitable algorithm for their specific 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 detection of moving object in real-world have fuelled 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: Frame Difference Based techniques, Deep Learning Based techniques, Optical Flow Based technique and Background Modelling 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 centre 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 detection of moving object in dynamic, scenarios of real world 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: Optical Flow Based techniques, Background Modelling Based techniques, Deep Learning Based techniques and Frame Difference Based techniques. Let's illustrate these categories with examples**:**

Background Modelling Based Techniques:

Example: The Gaussian Mixture Model (GMM) is a classic background modelling method. It characterizes the background of a video scene as a Gaussian distribution's mixture. 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 analysing 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: Detection of moving object has been revolutionized by Convolution neural Networks. Models like Faster R-CNN and YOLO (You Only Look Once) 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.

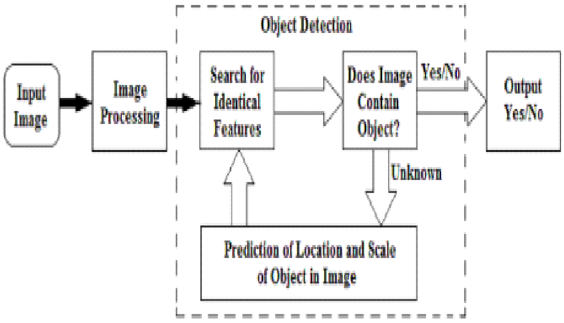

Fig: Object detection

General methodology

The object detection (OD) system primarily operates in two key phases: the learning phase and the testing phase as explained earlier in figure 2. The learning phase helps to determine objects within the input image by the system. This phase is divided into two subcategories: part-based or patch-based methods, a learning scheme is defined such that the system is trained in a process called learning through training. Learning through validation on the other hand learning through validation is a process in which the outcome of the learning through training process is validated or tested to ensure it can perform the intended task. At this stage, an object template block employs the obtained data to depict the objects by means of histogram or random forest data representations. This is done to remove the need for training since validation has been conducted before in the learning through validation phase. In the proposed method, once the input image has been pre-processed, template matching is directly used for extracting the features of an object in the image The testing phase in its turn is to decide whether or not an object exists in the image and if it does, which class it belongs to. Potential objects in the image are searched using a number of search methodologies including the sliding window methodology. Taking into account what has been discovered through the established searching mechanisms, a decision is then made on the class of the object.

Classification of OD mechanisms

This section classifies different forms of OD mechanisms depending on methods such as search techniques, features, classification, and features-based template matching mechanisms. It is for this reason that we have been able to categorise OD types which include: sliding window-based, contour-based, graph-based, fuzzy-based, context-based and others. In the following section, we summarize works done by several authors on object detection.

Sliding Window Object detection has received a lot of attention as its pioneered one of the simplest and essential approaches for detecting objects on images & videos. Starting from the window approach, the sliding window technique involves moving a window across the image or the entire scene in order to find the object of interest. However, this approach causes several difficulties with accurate positioning, especially in real-time applications, because of a large amount of computational work. It is much more critical in cases where object detection is performed, and is followed with recognition+.

Different approaches for realisation of context-based object detection in paintings were investigated by Bergère et al., namely the gradient method and the context detection method. The gradient method encodes the spatial context into a gradient pointing towards the object, while the context detection method uses a sliding window to select possible regions of interest in an image. Although the gradient method is derived under some assumptions and has the potential to create timing issues, particularly regarding the single-object detection, the sliding window approach has its timing issues, owing to its search approach.

This has been a major shortcoming when using the sliding window technique since localization has been inaccurate. Referring to nearby detection responses, Service et al. (2011) pointed out that elimination of the requirement for spatial grouping of nearby responses contributes to enhancing localization precision. Their findings highlighted three primary objectives in object detection: Such metrics as high recall rate, high precision and accuracy of localisation. They pointed out that pruning false positives on the basis of spatial clustering could lead to growth of localization variance.

Generally, the sliding window approach implies the identification of the size of the aforesaid search window. In order to overcome this issue, Comanche et al. (2013) predicted an approach to reduce the size of the sliding window in terms of its step size in runtime. This advancement enhances the technique's suitability for real-time applications. They proved that using their approach the Viola-Jones object detection framework is accelerated and the speed is increased by 2.03 frames per second without losing the accuracy. However, questions arose about space, again.

Divola (2012) identified two factors influencing the sliding window technique's performance: context and subcategories. Contextualization enables the approach to be directed to specific areas of an image where objects may be located instead of searching the entire image. Sub-categories mean cutting the much-diversified training data into smaller classes which makes the classification easier. Nevertheless, while only these factors were laid down, there are others like contour analysis which can greatly affect performance.

Shave brought forward a way to reduce the no detection misdetections while on the other hand increase the amount of space between each of the boxes in the sliding window-based approach. Their approach entails applying a patch to estimate the coordinates of an object in the search region's bounding box. To pile on the speed even more, they pre-charged this iterator with a decision tree, which has only simple binary tests at each node. However, while this system is said to work well on other image types, issues like occlusion are still unaddressed.

Future Challenging's:

The future challenges in object detection within computer vision:

Real-Time Object Detection at Scale:

Challenge: Detecting objects in real-time, particularly in high-resolution videos, requires significant computational power. Balancing the need for real-time performance with maintaining high accuracy is a major challenge.

Solution: To tackle this issue, the development of optimized network architectures, leveraging hardware accelerators such as GPUs and TPUs, and employing model quantization techniques offer potential solutions for enhancing efficiency and performance.

Handling Extremely Large Datasets:

Challenge: The expansion of deep learning and the increasing availability of large datasets create significant demands on memory, storage, and computational resources needed to train object detection models.

Solution: To address these challenges, approaches such as distributed training, data augmentation, and the selection of representative data subsets are essential for efficiently scaling object detection systems.

Robustness to Adverse Conditions:

Challenge: Real-world objects often face challenging environments, including poor lighting, harsh weather, occlusions, and cluttered backgrounds. Object detection systems must be capable of handling these conditions effectively.

Solution: Enhancing robustness can be achieved through data augmentation with synthetic conditions simulating these challenges, utilizing sensor fusion (such as combining RGB and infrared data), and implementing advanced pre-processing methods

Generalization Across Object Categories:

Challenge: Object detection models must be able to generalize across a broad range of object categories, including those that are rare or not previously encountered, even when only a small amount of labelled data is available.

Solution: Techniques such as few-shot learning, transfer learning, and meta-learning can help improve the model's capacity to adapt to new object categories with minimal annotated data.

Reducing Annotation Effort:

Challenge: **Challenge:** Annotating objects in images manually is both costly and time-consuming. The challenge is to minimize annotation efforts without compromising the accuracy of object detection.

Solution: Approaches like weakly supervised learning, active learning, and self-supervised learning can help reduce the reliance on extensive manual annotation.

Interpretability and Explainability:

Challenge: In safety-critical applications, it is essential to understand the reasoning behind an object detector's decisions. However, black-box models often lack transparency.

Solution: Creating models that offer interpretable outputs, such as attention mechanisms and visual explanations, enables users to comprehend the reasoning behind model decisions.

Privacy and Ethical Concerns:

Challenge: Object detection in public spaces brings up ethical issues related to privacy, surveillance, and the potential for misuse.

Solution: It is crucial to find a balance between the advantages and ethical concerns by implementing appropriate legislation, privacy-preserving techniques (such as federated learning), and transparent deployment practices.

Handling 3D Objects:

Challenge: In fields such as robotics and augmented reality, detecting and recognizing objects in three-dimensional space introduces unique challenges.

Solution: To address these challenges, combining 2D and 3D detection techniques, utilizing depth information (such as from LiDAR or depth cameras), and advancing 3D object detection networks are essential.

Adversarial Attacks:

Challenge: Adversarial attacks have the potential to deceive object detectors into producing incorrect predictions. Ensuring robustness against these attacks is essential.

Solution: Approaches such as adversarial training, evaluating model robustness, and integrating adversarial detection mechanisms are key to countering these threats.

Energy-Efficient Object Detection:

Challenge: In environments with limited resources, energy-efficient object detection is crucial. The challenge lies in optimizing models to consume less power without compromising accuracy.

Solution: Techniques such as model pruning, quantization, and hardware optimizations designed for energy-efficient inference are essential solutions.

Cross-Modal Object Detection:

Challenge: Combining data from various sensors (such as cameras and LiDAR) to enhance object detection in applications like autonomous vehicles demands effective cross-modal techniques.

Solution: Strategies for multi-modal fusion, sensor calibration, and cross-modal object detection models are key to successfully integrating sensor data.

Overcoming these complex challenges requires collaboration across multiple disciplines, including computer vision, hardware engineering, ethics, and policy-making. Additionally, continuous dialogue on the responsible deployment of AI is crucial to ensure that object detection technologies align with societal values and ethical standards.

Recommendation:

To push the boundaries of object detection, a comprehensive strategy is essential. Key to this is substantial investment in research and development, fostering innovation through collaboration between academic institutions, industry leaders, and government organizations. Ensuring open access to a wide variety of datasets and standardized evaluation metrics is critical for fair comparison and benchmarking. Efficiency is a primary concern, requiring the design and optimization of model architectures capable of real-time detection across different hardware platforms, along with exploring methods such as quantization and neural architecture search to reduce computational needs. Additionally, improving data efficiency through approaches like transfer learning and active learning is vital, allowing object detectors to make the most of limited annotated data. Research should focus on enhancing robustness, ensuring that models can handle diverse real-world conditions and defend against adversarial attacks. The development of interpretable models and ethical frameworks is crucial to ensure transparency and responsible AI practices. In contexts like surveillance and data-sharing, privacy-preserving methods such as federated learning should be prioritized. Cross-disciplinary collaboration between computer vision and sensor technology experts is key for applications in autonomous vehicles. Energy efficiency must also be addressed, with hardware optimizations and public awareness initiatives playing crucial roles. Establishing regulations, standards, and continuous evaluation platforms is necessary to monitor progress and promote responsible deployment. Ultimately, encouraging diversity and fostering global cooperation will lead to more inclusive and beneficial object detection solutions worldwide.

Below are several recommendations and strategies aimed at overcoming the challenges in object detection and driving progress in the field:

Research and Development:

- Dedicate resources to continuous research efforts to drive innovation in object detection methods.
- Foster collaboration among academic institutions, industry leaders, and government bodies to enhance research initiatives.

Open Access to Datasets and Benchmarks:

- Support the creation and sharing of diverse, comprehensive datasets for evaluating and benchmarking object detection algorithms.

- Advocate for the development of standardized evaluation metrics to ensure fair comparisons across different approaches.

Efficient Model Architectures:

- Focus on the design and optimization of model architectures that can achieve real-time object detection across a range of hardware platforms.
- Investigate model quantization and neural architecture search methods to minimize computational demands.

Data Efficiency and Transfer Learning:

- Create techniques that allow object detectors to learn effectively from a limited amount of annotated data using transfer learning, few-shot learning, and semi-supervised learning approaches.
- Emphasize active learning strategies to selectively choose the most informative samples for annotation.

Robustness and Adversarial Défense:

- Focus on advancing research into defence mechanisms that enhance the resilience of object detection systems against adversarial attacks.
- Regularly assess models in diverse and challenging real-world scenarios to ensure their reliability.

Interpretable Models:

- Encourage research into interpretable deep learning models that offer transparency and insights into how decisions are made.
- Develop visualization tools and methods to help users better understand and interpret model outputs.

Ethical Considerations:

- Develop ethical guidelines and codes of conduct for the creation and deployment of object detection systems.
- Promote responsible AI practices and take into account the broader societal effects of these technologies.

Privacy-Preserving Techniques:

- Investigate privacy-preserving strategies, such as federated learning and secure multi-party computation, to safeguard individual privacy in surveillance and data-sharing contexts.

Cross-Modal Integration:

- Encourage collaboration between computer vision experts and sensor technology specialists to enhance cross-modal object detection, especially for applications like autonomous vehicles.

Energy Efficiency:

- Explore hardware-specific optimizations and model compression techniques to enable energy-efficient object detection on edge devices and mobile platforms.

Public Awareness and Education:

- Raise public awareness about the capabilities, limitations, and ethical concerns associated with object detection technologies.
- Promote public engagement and involve relevant stakeholders in shaping regulations and policies.

Regulations and Standards:

- Advocate for the establishment of regulations and standards governing the application of object detection in sensitive areas, ensuring ethical and responsible usage.

Continuous Evaluation and Benchmarking:

- Develop platforms for ongoing assessment and comparison of object detection models to monitor progress and pinpoint areas for improvement.

Diverse Representation:

- Foster diversity and inclusivity in object detection research to ensure that systems are effective across all demographic groups.

International Collaboration:

- Promote global collaboration and knowledge exchange among researchers, organizations, and governments to tackle shared challenges and advance responsible AI.

By implementing these strategies, object detection can evolve, providing more accurate, efficient, and ethical solutions that address societal needs and improve a variety of applications.

Conclusion

In conclusion, object detection remains a rapidly evolving and crucial area within computer vision, offering a variety of applications that influence everyday life. As we explore the complexities of real-world environments, it becomes clear that both challenges and opportunities exist in abundance. This review has highlighted the wide spectrum of techniques in moving object detection, ranging from background modelling and frame differencing to optical flow analysis and deep learning methods. It has emphasized the significance of these approaches in fields such as surveillance, robotics, autonomous driving, and augmented reality.

Looking ahead, continued investment in research and development is critical to driving innovation and collaboration among academia, industry, and policymakers. The availability of open datasets and the adoption of standardized evaluation metrics will facilitate fair comparisons of detection algorithms. It is also vital that we prioritize efficiency and data effectiveness, focusing on optimizing model architectures, as well as exploring transfer learning and active learning techniques.

Additionally, addressing the robustness of models against adversarial attacks and enhancing their interpretability to build user trust will be crucial. Ethical considerations should not be overlooked, and efforts must be made to ensure responsible AI development and the implementation of privacy-protecting technologies.

The future of object detection will also benefit from advancements in cross-modal integration, energy efficiency, public engagement, and international cooperation. The establishment of regulations and standards to govern sensitive use cases, along with ongoing evaluation platforms, is essential for ensuring responsible application and deployment.

Ultimately, the future of object detection offers immense potential for technological innovation and societal impact. By addressing current challenges and capitalizing on emerging opportunities, we can continue advancing this field and creating more accurate, efficient, and ethically responsible solutions that contribute to the betterment of society.

References

- *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)*
- *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)*
- *Nicholson, J., Kulyukin, V., Coster, D.: ShopTalk: independent blind shopping through verbal route directions and barcode scans. Open Rehabil. J. 2, 11–23 (2009)*
- *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)*
- *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)*
- *Kutiyanawala, A., Kulyukin, V.: An eyes-free vision-based UPC and MSI barcode localization and decoding algorithm for mobile phones. In: Envision 2010 (2010)*
- *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)*
- *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)*
- *Semacode Corporation. http://semacode.com/about/*
- *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)*
- *Lowe, D.G.: Object recognition from local scale-invariant features. In: The Seventh IEEE International*

Conference on Computer Vision, Kerkyra, pp. 1150–1157 (1999)

- *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)*
- *Apple Inc.* *http://www.apple.com/iphone/*
- *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*
- *Leibs, A.: Top 10 iPhone Apps for the Visually Impaired. 22 August 2012*
- *Hub, A., Hartter, T., Ertl, T.: Interactive tracking of movable objects for the blind on the basis of environment models and perception-oriented object recognition methods. In: Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility, Portland, Oregon, USA, pp. 111–118 (2006)*
- *Hub, A., Hartter, T., Ertl, T.: Interactive localization and recognition of objects for the blind. In: Northridge Center on Disabilities' 21st Annual International Technology and Persons with Disabilities Conference. California State University (2006)*
- *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)*
- *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)*
- *Blauert, J.: Spatial Hearing: The Psychophysics of Human Sound Localization. MIT Press, Cambridge (1997)*
- *Lowe, D.G.: Distinctive image features from scale-invariant key points. Int. J. Comput. Vis. 60, 91–110 (2004)*
- *Bakken, T.: An evaluation of the SIFT algorithm for CBIR. Telenor R&I N 30/2007, 2007*
- *Ramisa, A., Vasudevan, S., Advert, D., Toledo, R., Mantras, 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)

- *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. Springer, Berlin/Heidelberg (2012)*
- *Mikolajczyk, K., Schmid, C.: Scale and affine invariant interest point detectors. Int. J. Comput. Vis. 60, 63–86 (2004)*
- *Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Trans. Pattern Anal. Mach. Intell. 27, 1615–1630 (2005)*
- *Pavlidis, T.: An evaluation of the Scale Invariant Feature Transform (SIFT). (25 August 2008)*
- *Ancuti, C., Bekaert, P.: Sift-cch: increasing the sift distinctness by colour co-occurrence histograms. In: The Fifth International Symposium on Parallel and Distributed Processing and Applications (ISPA07), pp. 130–135 (2007)*
- *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)*
- *Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-up robust features (SURF). Comput. Vis. Image Underst. 110, 346–359 (2008)*
- *Hasanuzzaman, F.M., Yang, X., Tian, Y.: Robust and effective component-based banknote recognition by SURF features. In: Proc. WOCC, pp. 1–6 (2011)*
- *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)*
- *Winlock, T., Christiansen, E., Belongie, S.: Toward real-time grocery detection for the visually impaired. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 49–56 (2010)*
- *Amazon mechanical turk. http://www.mturk.com/*
- *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)*
- *Swain, M., Ballard, D.: Color indexing. Int. J. Comput. Vis. 7, 11–32 (1991)*