Designing Transparent and Autonomous Intelligent Vision Systems

Joanna Isabelle Olszewska
School of Engineering and Computing, University of West Scotland, U.K.

Keywords: Intelligent Vision Systems, Autonomous Systems, Vision Agents, Agent-oriented Software Engineering, Explainable Artificial Intelligence (XAI), Transparency and Ethical Issues.

Abstract: To process vast amounts of visual data such as images, videos, etc. in an automatic and computationally efficient way, intelligent vision systems have been developed over the last three decades. However, with the increasing development of complex technologies like companion robots which require advanced machine vision capabilities and, on the other hand, the growing attention to data security and privacy, the design of intelligent vision systems faces new challenges such as autonomy and transparency. Hence, in this paper, we propose to define the main requirements for the new generation of intelligent vision systems (IVS) we demonstrated in a prototype.

1 INTRODUCTION

With the omnipresence of digital data in our Society, and in particular visual data (Olszewska, 2018) from smartphone pictures to television video streams, from m-health services to social media apps, from street surveillance cameras to airport e-gates, from drones to autonomous underwater vehicles (AUVs), intelligent vision systems (IVS) are needed to automatize the processing of these visual data.

Indeed, intelligent vision systems are a set of interconnected hardware and/or software components which take digital image(s) as input data and process them by means of methods ranging from low- to high-level techniques/algorithms in order to extract meaningful information, which could be structured and organized into knowledge, and aid to the automatic understanding of the gathered visual data (Fig.1).

The input image to be processed by an IVS could be a single, still picture, a set of pictures, or a dynamic stream of frames (Sabour et al., 2008). The image(s) could be gathered from an online/offline database (Rahbi et al., 2016) or acquired live by a single vision sensor or multiple ones (Bianchi and Rillo, 1996), each sensor being either static or dynamic (Ishiguro et al., 1993).

The IVS output is the information obtained after processing the input visual data. The resulting information could present any degree of modality, i.e. could be of a semantic type (degree 0) such as a tag/label or a text file, of a visual type itself, e.g. a picture or a region of a picture (degree 1), or of a video type (degree 2). This information could also consist in further processed results such as computed trajectories (Ukita and Matsuyama, 2002), workflows (Sardis et al., 2010), etc. It could produce knowledge (Reichard, 2004) and/or provide further understanding of the visual data (Li et al., 2009).

In order to process the input visual data, a (series of) method(s) is implemented within the intelligent vision system. It could consist of low-level techniques of image processing such as thresholding, morphological operations, edge detection, texture region detection (Arbuckle and Beetz, 1999), etc. Mid-level techniques include computer vision methods such as interest point descriptors, active contours (Olszewska, 2015), etc. High-level techniques involve artificial intelligence (AI), and in particular, machine learning approaches which could be based on symbolism (e.g. logic rules) (Olszewska, 2017), analogism (e.g. Support Vector Machine (SVM)) (Prakash et al., 2012), probability (e.g. Bayesian rule) (Hou et al., 2014), evolutionarism (e.g. genetic algorithms) (Nestinger and Cheng, 2010), or connectivism (e.g. neural networks) (Jeon et al., 2018).

The design of such complex systems necessitates a careful requirement analysis (Rash et al., 2006), not only in terms of performance targets of the vision algorithm(s), but also in terms of software and system requirements.

Olszewska, J.
Designing Transparent and Autonomous Intelligent Vision Systems.
DOI: 10.5220/0007585208500856
In *Proceedings of the 11th International Conference on Agents and Artificial Intelligence (ICAART 2019)*, pages 850-856
ISBN: 978-989-758-350-6

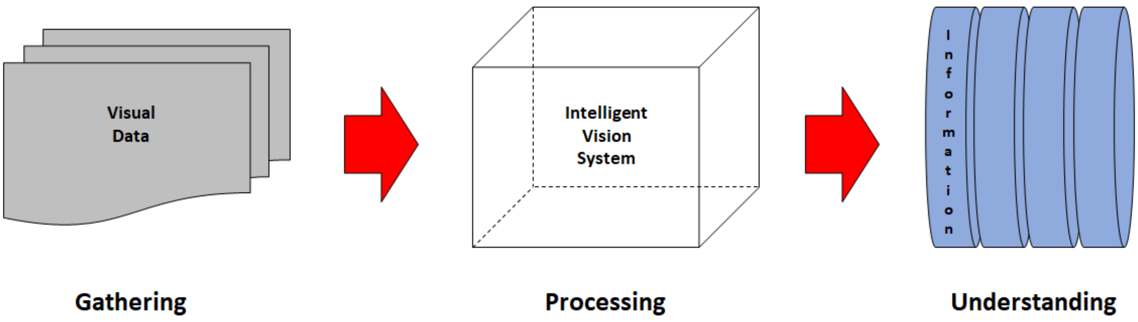

Figure 1: Intelligent vision system (IVS) overview.

Indeed, in this paper, we present the requirements for intelligent vision systems.

IVS can be considered either as an intelligent vision software or an embodiment of a vision process.

Therefore, the proposed set of requirements provides the backbone for the ethical and dependable design of such complex vision systems.

The main contributions of this work are, on one hand, the identification of the requirements for the new generation of intelligent vision systems and, on the other hand, the definition of the main concepts such as autonomy and transparency in context of intelligent vision systems as well as the elucidation of the intelligent vision system notion itself and its related design pattern.

The paper is structured as follows. In Section 2, we present the intelligent vision system requirements and their related definitions. The proposed design method has been successfully prototyped as reported and discussed in Section 3. Conclusions are drawn up in Section 4.

2 REQUIREMENTS

The proposed IVS requirements are intelligent vision system's reliability, security, autonomy, and transparency, as described in Sections 2.1-2.4, respectively.

2.1 Reliable

The prime focus of the IVS design has been the efficiency of such systems in order to develop reliable solutions.

Low- and mid-level IVS performance are assessed using one or more metrics quantifying shape fidelity (Correia and Pereira, 2003), shape accuracy (Gelasca and Ebrahimi, 2009), shape temporal coherence (Erdem and Sankur, 2000), connectivity, and compactness (Goumeidane and Khamadja, 2010).

However, there is no single definition of these measures. In particular, the segmentation error could be defined in several ways, e.g. by calculating:

- the Pratt's Figure of Merit (FOM) (Pratt et al., 1978) which evaluates the edge location accuracy via the displacement of detected edge points from an ideal edge:

$$FOM = \frac{1}{max(I_I, I_A)} \sum_{i=1}^{I_A} \frac{1}{1+\delta\, d^2(i)}, \tag{1}$$

  with I_I, the number of ideal edge points, I_A, the number of actual edge points, d, the Euclidean distance between the actual points and the ideal edge points, and δ, a scaling constant (of usual value $\delta = 1/9$). The value of FOM falls between 0 and 1, and the larger, the better.

- the overlap (OL) between the segmented region A and the ground truth region A_G (Saeed and Dugelay, 2010):

$$OL = \frac{2(A \cap) A_G}{A + A_G} * 100, \tag{2}$$

  where $OL \in [0, 100]$, and the larger, the better.

- the segmentation error (SE) for edge and region-based methods (Saeed and Dugelay, 2010):

$$SE = \frac{A}{2 * A_G} * 100, \tag{3}$$

  where $SE \in [0, 100]$, and the smaller, the better.

- the spatial accuracy or MPEG-4 Segmentation error Se (Muller-Schneiders et al., 2005):

$$Se = \frac{\sum_{k=1}^{fn} d_{fn}^k + \sum_{l=1}^{fp} d_{fp}^l}{card(M_r)}, \tag{4}$$

  where $card(M_r)$ is the number of pixels of the reference mask; *fn*, the number of false negative pixels; *fp*, the number of false positive pixels; d_{fn}^k, the distance of the k^{th} false negative pixel to the reference pixel; and d_{fp}^l, the distance of the l^{th} false

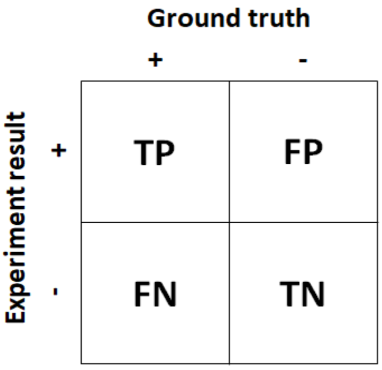

Figure 2: Representation of the true positive rate (TP), the false positive rate (FP), the false negative rate (FN), and the true negative rate (TN), respectively, in a confusion matrix.

positive pixel to the reference pixel. The value of Se is in the range of $[0, \infty[$, and the smaller, the better.

The above-mentioned measures (Eqs. 1-4) could lead to further assessment of IVS, e.g. by plotting detected/tracked object's actual trajectory vs ground truth one (Leibe et al., 2008).

The effectiveness of mid- and high-level IVS is mainly measured using the following set of metrics (Estrada and Jepson, 2005):

$$precision\ (P) = \frac{TP}{TP+FP}, \tag{5}$$

$$detection\ rate\ (DR)\ or\ \ (R) = \frac{TP}{TP+FN}, \tag{6}$$

$$false\ \ detection\ rate\ (FAR) = \frac{FP}{FP+TP}, \tag{7}$$

$$accuracy\ (ACC) = \frac{TP+TN}{TP+TN+FP+FN}, \tag{8}$$

where TP is the True positive rate, FP is the False Positive rate, FN is the False Negative rate, and TN is the True Negative rate (see Fig. 2).

It is worth noting the detection rate (DR) is also sometimes called recall (R), sensitivity, or hit rate.

Another common metric is the F1 score which is the harmonic mean of the precision and recall and which could be used when a balance between precision and recall is needed and when the class distribution is uneven (i.e. high $TN+FP$). F1 score is defined as follows:

$$F1\ score\ (F1) = 2\frac{P*R}{P+R}. \tag{9}$$

To represent and evaluate IVS efficiency, the standard measures (Eqs. 5-8) could be used to compose a confusion matrix such as in Fig. 2 and/or a precision-recall curve. The later one is useful in the case the classes are very imbalanced and shows the trade-off between precision and recall for different thresholds. It is worth noting a high area under the precision-recall curve is aimed, since it represents both high recall (i.e. low FN) and high precision (i.e. low FP). Indeed, high scores for both precision and recall show that the system is returning accurate results (high precision) as well as returning a majority of all positive results (high recall), i.e. the system returns many results, with most of the results labeled correctly. Incidentally, a system with high recall but low precision returns many results, but most of its predicted labels are incorrect when compared to the training labels. A system with high precision but low recall returns very few results, but most of its predicted labels are correct when compared to the training labels.

On the other hand, programming paradigms and implementation language choices have an impact on the IVS design, with C++/OpenCV and MatLab the most common languages adopted for IVS (Nestinger and Cheng, 2010). Moreover, the development of reliable IVS is a Test-Driven Development (TDD) that can be assessed using software engineering indicators (Pressman, 2010) such as function-based metrics, specification quality metrics, architectural design metrics, class-oriented metrics, component-level design metrics (including cohesion and coupling metrics), operation-oriented metrics, user-interface design metrics, etc. Hence, IVS code complexity, maintainability, and quality (Reichardt et al., 2018) as well as IVS computational speed (e.g. computational time vs image resolution/size) are crucial to be analysed in the IVS design phase. .

Furthermore, IVS robustness and fault tolerance (Asmare et al., 2012) along with its portability and interoperability (Bayat et al., 2016) should also be considered in the IVS design phase, as IVS tend to be deployed not only on desktops/laptops, but also on smartphones, robots, industrial equipment, etc.

2.2 Secure

In today's society, ensuring the security of cyber-physical systems is a main challenge (Escudero et al., 2018), (Burzio et al., 2018). Therefore, in addition of correctness and robustness (Reichard, 2004), IVS design should integrate the cybersecurity element (Russell et al., 2015). For example, the communication between the visual data acquisition set-up and the machine processing them should be secured (Leonard et al., 2017).

Furthermore, the collected visual data which are

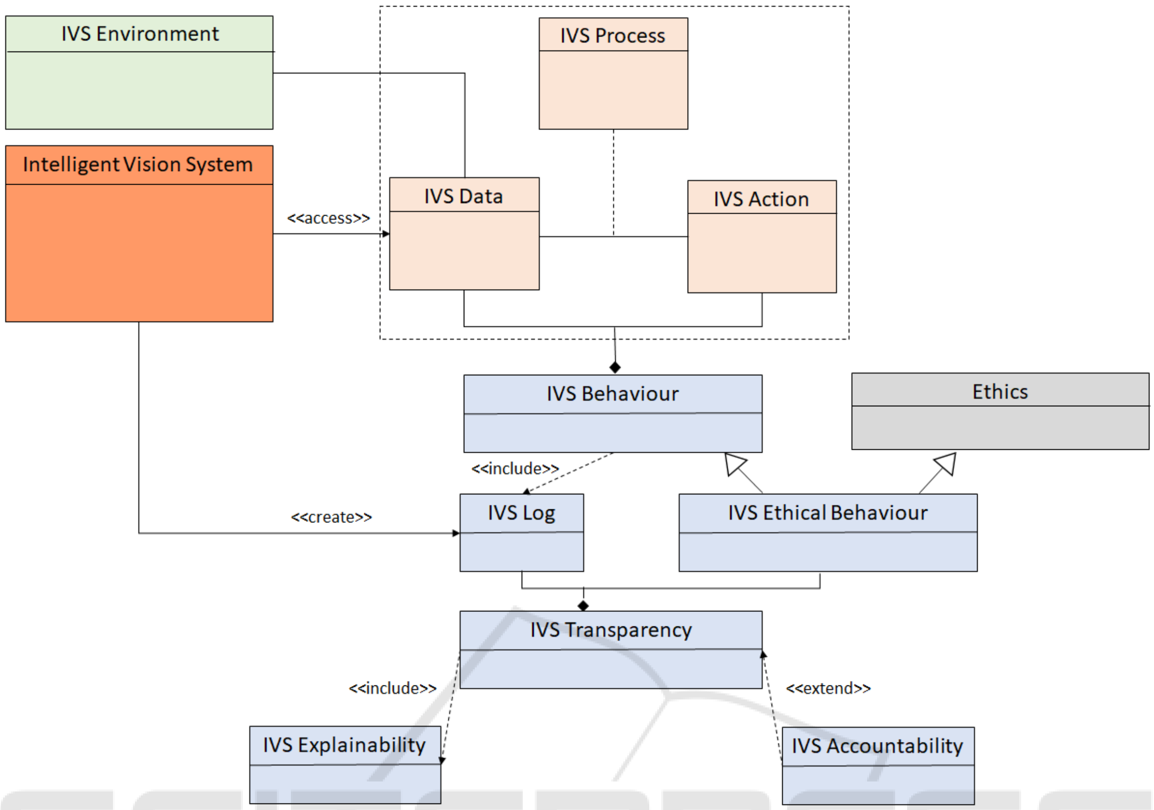

Figure 3: UML Class Diagram of the Intelligent Vision System.

processed by IVS as well as the produced information and knowledge should respect data privacy and in particular comply with GDPR legislation (General Data Protection Regulation (GDPR), 2018).

Hence, the design of such correct, robust, and secure IVS ensures its dependability (Meyer, 2006).

2.3 Autonomous

Autonomous systems (AS) are systems that decide themselves what to do and when to do it (Fisher et al., 2013). Their features could include failure diagnosis, self-awareness, biomimetic, automated reasoning or knowledge-inspired transactions (Olszewska and Allison, 2018).

IVS autonomy is the system's capacity of managing itself using artificial intelligence (AI) method(s) to produce the intended goal(s), i.e. to process and provide the expected information, to build the relevant knowledge, and/or to automatically analyse/understand its environment. Depending of its level of autonomy (i.e. not autonomous, semi-autonomous, autonomous), the system will have a various degree of interactions with other intelligent agents (Calzado et al., 2018) through manned or remote control operations (Zhang et al., 2017).

Figure 3 presents a possible pattern for such systems, whatever their classification (Franklin and Graesser, 1996), architecture (Fiorini et al., 2017), or environment (Osorio et al., 2010).

2.4 Transparent

Transparency in AS is the property which makes possible to discover how and why the system made a particular decision or acted the way it did (Chatila et al., 2017), taking into account its environment (Lakhmani et al., 2016).

IVS transparency is the system's capacity of its goals, its situational constraints, its input/output data, its decision criteria, its internal structure, its assumptions about the external reality, its actions, and its interactions being understood by the relevant stakeholders. Hence, depending of the level of transparency, the system could be transparent to the system's users, commanders, regulators, and/or investigators.

Therefore, an IVS system should adopt the design pattern proposed in Fig. 3. Moreover, designing transparent IVS implies the choice of intelligent techniques, like the machine learning (ML) methods, not only in terms of efficiency but also in terms of transparency. As per (Bostrom and Yudkowsky, 2014), the most transparent ML techniques are logic based and the less transparent ones involve neural networks. Indeed, with explainable AI (XAI) (Ha et al., 2018) being in its infancy, logic-based approaches (Olszewska and McCluskey, 2011) have the most well-established procedures for system verification (Brutzman et al., 2012). In addition, ML training databases should avoid biases (Skirpan and Yeh, 2017).

3 PROTOTYPE

The proposed requirements have been demonstrated on a IVS prototype aiming to detect and count people in indoor environments (Fig. 4). The system can be used by a university to monitor the use of computer labs during assignment deadline times. This can help a university to determine whether the computer labs should be open for longer during these periods. Another benefit of the developed application is that it can be used for safety purposes to provide an accurate number of people inside a building, aiding the evacuation process e.g. in the event of a fire.

The IVS prototype has been run on a computer with Intel Core i5-2400 CPU, 3.10 GHz, 12Gb RAM, 64-bit Windows 7 Enterprise OS, using MatLab language.

As per Section 2, the IVS reliability has been assessed in terms of recall ($R = 80\%$) and precision ($P = 91\%$). The computational speed rate is 14 fps.

The IVS security has been analysed and the data privacy has been ensured in order people appearing within the video will remain anonymous.

The IVS system has been designed in order to work in an autonomous way, performing the tasks as follows:

- Detecting people - the software is able to detect the people in the video automatically.
- Tracking people - the software automatically tracks people in the video, once they are detected.
- Counting people - the software is able to automatically count the people detected in the video.

The IVS prototype has been designed in order to be transparent to both users and experts. Indeed, from the user's point of view, the input for this software is the video-recorded stream that the software processes, and the software output is in the form of two video players that show simultaneously the annotated input and the extracted information all along the software execution (Fig. 4). From the expert's point of view, the software process does not involve deep learning and is rather based on the main intelligent vision techniques such as background subtraction and foreground detection by thresholding, blob analysis by applying morphological mathematic operations, tracking using a Kalman filter (KF), and a counter algorithm to count the people detected; all these techniques being formally verifiable.

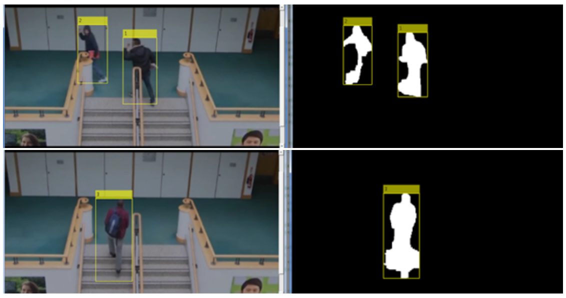

Figure 4: Example of the intelligent vision system prototype in action.

4 CONCLUSIONS

In this work, we proposed a set of requirements and definitions to aid the design of (near-)future intelligent vision systems (IVS). Indeed, intelligent vision systems have been developed for 30 years, with a main focus on their efficiency. Nowadays, IVS presence is ever growing in our daily life. Hence, their design should not only take into account their reliability in terms of accuracy and robustness, but also encompass concepts such as transparency, autonomy, security, and privacy. Experiments in real-world context have displayed the effectiveness and usefulness of our approach well-suited for applications within image/video processing, pattern recognition, computer vision, and machine vision domains.

ACKNOWLEDGEMENTS

The author would like to thank her students, R. Dobie and J. Webb for their involvement in the prototype.

REFERENCES

Arbuckle, T. and Beetz, M. (1999). Controlling image processing: Providing extensible, run-time configurable functionality on autonomous robots. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 787–792.

Asmare, E., Gopalan, A., Sloman, M., Dulay, N., and Lupu, E. (2012). Self-management framework for mobile autonomous systems. *Journal of Network and Systems Management*, 20(2):244–275.

Bayat, B., Bermejo-Alonso, J., Carbonera, J. L., Facchinetti, T., Fiorini, S., Goncalves, P., Jorge, V. A. M., Habib, M., Khamis, A., Melo, K., Nguyen, B., Olszewska, J. I., Paull, L., Prestes, E., Ragavan, V., Saeedi, S., Sanz, R., Seto, M., Spencer, B., Vosughi, A., and Li, H. (2016). Requirements for building an ontology for autonomous robots. *Industrial Robot: An International Journal*, 43(5):469–480.

Bianchi, R. A. C. and Rillo, A. H. R. C. (1996). A purposive computer vision system: A multi-agent approach. In *Proceedings of the IEEE Workshop on Cybernetic Vision*, pages 225–230.

Bostrom, N. and Yudkowsky, E. (2014). The ethics of artificial intelligence. *The Cambridge Handbook of Artificial Intelligence*.

Brutzman, D., McGhee, R., and Davis, D. (2012). An implemented universal mission controller with run time ethics checking for autonomous unmanned vehicles — a uuv example. In *Proceedings of the IEEE International Conference on IEEE/OES Autonomous Underwater Vehicles*, pages 1–8.

Burzio, G., Cordella, G. F., Colajanni, M., Marchetti, M., and Stabili, D. (2018). Cybersecurity of connected autonomous vehicles : A ranking based approach. In *Proceedings of the IEEE International Conference of Electrical and Electronic Technologies for Automotive*, pages 1–6.

Calzado, J., Lindsay, A., Chen, C., Samuels, G., and Olszewska, J. I. (2018). SAMI: Interactive, Multi-Sense Robot Architecture. In *Proceedings of the IEEE International Conference on Intelligent Engineering Systems*, pages 317–322.

Chatila, R., Firth-Butterflied, K., Havens, J. C., and Karachalios, K. (2017). The IEEE global initiative for ethical considerations in artificial intelligence and autonomous systems. *IEEE Robotics and Automation Magazine*, 24(1):110.

Correia, P. L. and Pereira, F. (2003). Objective evaluation of video segmentation quality. *IEEE Transactions on Image Processing*, 12(2):186–200.

Erdem, C. E. and Sankur, B. (2000). Performance evaluation metrics for object-based video segmentation. In *Proceedings of European Signal Processing Conference (EUSPICO)*, pages 1–4.

Escudero, C., Sicard, F., and Zamai, E. (2018). Process-aware model based idss for industrial control systems cybersecurity: Approaches, limits and further research. In *Proceedings of the IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 605–612.

Estrada, F. J. and Jepson, A. D. (2005). Quantitative evaluation of a novel image segmentation algorithm. In *Proceedings of the Joint IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1132–1139.

Fiorini, S. R., Bermejo-Alonso, J., Goncalves, P., de Freitas, E. P., Alarcos, A. O., Olszewska, J. I., Prestes, E., Schlenoff, C., Ragavan, S. V., Redfield, S., Spencer, B., and Li, H. (2017). A suite of ontologies for robotics and automation. *IEEE Robotics and Automation Magazine*, 24(1):8–11.

Fisher, M., Dennis, L., and Webster, M. (2013). Verifying autonomous systems. *Communications of the ACM*, 5-6(9):84–93.

Franklin, S. and Graesser, A. (1996). Is it an agent, or just a program?: A taxonomy for autonomous agents. In *Proceedings of the Intelligent Agents III Agent Theories, Architectures, and Language, LNCS, Springer*, pages 21–35.

Gelasca, E. D. and Ebrahimi, T. (2009). On evaluating video object segmentation quality: A perceptually driven objective metric. *IEEE Journal of Selected Topics in Signal Processing*, 3(2):319–335.

General Data Protection Regulation (GDPR) (2018). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data.

Goumeidane, A. B. and Khamadja, M. (2010). Error measures for segmentation results: Evaluation on synthetic images. In *Proceedings of IEEE International Conference on Electronics, Circuits and Systems*, pages 158–161.

Ha, T., Lee, S., and Kim, S. (2018). Designing explainability of an artificial intelligence system. In *Proceedings of the ACM Conference on Technology, Mind, and Society*.

Hou, Y., Zhang, Y., Xue, F., Zheng, M., and Fan, R. (2014). The fusion method of improvement Bayes for soccer robot vision system. In *Proceedings of IEEE International Symposium on Computer, Consumer and Control*, pages 776–780.

Ishiguro, H., Kato, K., and Tsuji, S. (1993). Multiple vision agents navigating a mobile robot in a real world. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 772–777.

Jeon, H.-S., Kum, D.-S., and Jeong, W.-Y. (2018). Traffic scene prediction via deep learning: Introduction of multi-channel occupancy grid map as a scene representation. In *Proceedings of IEEE Intelligent Vehicles Symposium (IV)*, pages 1496–1501.

Lakhmani, S., Abich, J., Barber, D., and Chen, J. (2016). A proposed approach for determining the influence of multimodal robot-of-human transparency information on human-agent teams. In *Proceedings of the International Conference on Augmented Cognition*, pages 296–307.

Leibe, B., Schindler, K., Cornelis, N., and Gool, L. V. (2008). Coupled object detection and tracking from static cameras and moving vehicles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10):1683–1698.

Leonard, S. R., Allison, I. K., and Olszewska, J. I. (2017). Design and Test (D & T) of an in-flight entertainment system with camera modification. In *Proceedings of the IEEE International Conference on Intelligent Engineering Systems*, pages 151–156.

Li, L.-J., Socher, R., and Fei-Fei, L. (2009). Towards total scene understanding: Classification, annotation and segmentation in an automatic framework. In *Proceedings of the IEEE International Conference on Com-*

puter Vision and Pattern Recognition (CVPR), pages 2036–2043.

Meyer, B. (2006). Dependable software. In Kohlas, J., Meyer, B., and Schiper, A., editors, *Dependable Systems*, pages 1–33. Springer.

Muller-Schneiders, S., Jager, T., Loos, H. S., and Niem, W. (2005). Performance evaluation of a real time video surveillance system. In *Proceedings of the Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (PETS)*, pages 137–143.

Nestinger, S. S. and Cheng, H. H. (2010). Flexible vision. *IEEE Robotics and Automation Magazine*, 17(3):66–77.

Olszewska, J. (2017). Detecting hidden objects using efficient spatio-temporal knowledge representation. In *Revised Selected Papers of the International Conference on Agents and Artificial Intelligence. Lecture Notes in Computer Science*, volume 10162, pages 302–313.

Olszewska, J. I. (2015). Active contour based optical character recognition for automated scene understanding. *Neurocomputing*, 161:65–71.

Olszewska, J. I. (2018). Discover intelligent vision systems. In *DDD Scotland 2018*.

Olszewska, J. I. and Allison, I. K. (2018). ODYSSEY: Software Development Life Cycle Ontology. In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development*, pages 301–309.

Olszewska, J. I. and McCluskey, T. L. (2011). Ontology-coupled active contours for dynamic video scene understanding. In *Proceedings of the IEEE International Conference on Intelligent Engineering Systems*, pages 369–374.

Osorio, F., Wolf, D., Branco, K. C., and Pessin, G. (2010). Mobile robots design and implementation: From virtual simulation to real robots. In *Proceedings of IDMME 2010*, pages 1–6.

Prakash, J. S., Vignesh, K. A., Ashok, C., and Adithyan, R. (2012). Multi class support vector machines classifier for machine vision application. In *Proceedings of IEEE International Conference on Machine Vision and Image Processing*, pages 197–199.

Pratt, W., Faugeras, O. D., and Gagalowicz, A. (1978). Visual discrimination of stochastic texture fields. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(11):796–804.

Pressman, R. S. (2010). Product metrics. In *Software engineering: A practitioner's approach, 7th Ed.*, pages 613–643. McGraw-Hill.

Rahbi, M. S. A., Edirisinghe, E., and Fatima, S. (2016). Multi-agent based framework for person reidentification in video surveillance. In *Proceedings of the IEEE Future Technologies Conference*, pages 1349–1352.

Rash, J. L., Hinchey, M. G., Rouff, C. A., Gracanin, D., and Erickson, J. (2006). A requirements-based programming approach to developing a NASA autonomous ground control system. *Artificial Intelligence Review*, 25(4):285–297.

Reichard, K. M. (2004). Integrating self-health awareness in autonomous systems. *Robotics and Autonomous Systems*, 49(1-2):105–112.

Reichardt, M., Foehst, T., and Berns, K. (2018). Introducing finroc: A convenient real-time framework for robotics based on a systematic design approach.

Russell, S., Dewey, D., and Tegmark, M. (2015). Research priorities for robust and beneficial artificial intelligence. *AI Magazine*, 36(4):105–114.

Sabour, N. A., Faheem, H. M., and Khalifa, M. E. (2008). Multi-agent based framework for target tracking using a real time vision system. In *Proceedings of the IEEE International Conference on Computer Engineering and Systems*, pages 355–363.

Saeed, U. and Dugelay, J.-L. (2010). Combining edge detection and region segmentation for lip contour extraction. In *Proceedings of International Conference on Articulated Motion and Deformable Objects*, pages 11–20.

Sardis, E., Anagnostopoulos, V., and Varvarigou, T. (2010). Multi-agent based surveillance of workflows. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pages 419–422.

Skirpan, M. and Yeh, T. (2017). Designing a moral compass for the future of computer vision using speculative analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1368–1377.

Ukita, N. and Matsuyama, T. (2002). Real-time multi-target tracking by cooperative distributed active vision agents. In *Proceedings of the ACM International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 829–838.

Zhang, T., Li, Q., Zhang, C.-S., Liang, H.-W., Li, P., Wang, T.-M., Li, S., Zhu, Y.-L., and Wu, C. (2017). Current trends in the development of intelligent unmanned autonomous systems. *Frontiers of Information Technology and Electronic Engineering*, 18(1):68–85.