Solving operational tasks in the design of video surveillance systems

Nataliia Bahniuk[†], Taras Terletskyi[*,†], Oleh Kaidyk[†], Serhii Kostiuchko[†] and Inna Kondius[†]

Lutsk National Technical University, Lvivska Street 75, 43016 Lutsk, Ukraine

Abstract
The article discusses the problem of CCTV design, which consists in the constant change of recommended criteria for equipment selection caused by the continuous and rapid evolution of technologies. The results of the analysis of existing standards are presented and discrepancies are identified. This trend leads to the destabilization of clear quality criteria and the risk of unreasonable CCTV design. The paper highlights the methods and results of research into changes in the spatial resolution of images from a range of technical characteristics of video cameras based on the relevant theory, which is implemented analytically and confirmed by computer modeling using specialized software. The results of the study, obtained by calculation and modeling, indicated a general discrepancy in the data obtained, which does not exceed the permissible limits. The data presented in the study results will help designers make an informed choice of video cameras, which will reduce the risk of excessive design.

Keywords
Design, criteria, operational task, pixel density, matrix, image quality, scene, resolution.

1. Introduction

In the context of constantly growing global terrorist threats, characterized by a high degree of unpredictability and destructiveness, the relevance of video surveillance systems is becoming strategically important in the field of national and international security. This technological tool is transforming from a passive means of recording events into a proactive component of a comprehensive counterterrorism system, which is crucial for protecting society and its critical infrastructure, provided that the right approaches are taken to its design.

A video surveillance system is classified as a specialpurpose information television system. Internationally, it operates under the commonly accepted name Closed Circuit Television (CCTV), which indicates its nature as a closedcircuit television system accessible only to an authorized group of observers.

Since the tasks of these systems have expanded, it makes sense to consider the issue of solving potential operational tasks and criteria for their design.

2. Problem statement

The current problem with CCTV design is the lack of stable, universal criteria for selecting equipment, caused by the continuous and rapid evolution of technology. This was particularly noticeable during the transition from analog to digital systems. This dynamic creates the following dilemmas: destabilization of quality criteria (throughout the history of these systems, quality assessment criteria have constantly changed) and the risk of unjustified overengineering (the appropriate choice of video camera resolution).

[*] *AIT&AIS'2025: International Scientific Workshop on Applied Information Technologies and Artificial Intelligence Systems, December 18–19 2025, Chernivtsi, Ukraine*

[1*] Corresponding author.
[†] These authors contributed equally.
bagnyuknata@gmail.com (N. Bahniuk); t.terletskyi@lntu.edu.ua (T. Terletskyi); o.kaidyk@lntu.edu.ua (O. Kaidyk); s.kostiuchko@lutsk-ntu.com.ua (S. Kostiuchko); i.kondius@lntu.edu.ua (I. Kondius)
0000000271205455 (N. Bahniuk); 0000000241140734 (T. Terletskyi); 000000023620270X (O. Kaidyk); 0000000212626268 (S. Kostiuchko); 0000000334167506 (I. Kondius)

Designers face these questions when selecting equipment for a new system in accordance with modern requirements, which need to be addressed.

3. Formulation of the purpose of the article

The purpose of this work is to establish the influence of technical parameters of video cameras on the image quality of an object when solving existing types of operational CCTV tasks in accordance with the recommendations of existing standards. This will allow for a reasonable approach to the selection of video cameras and reduce the risk of overdesign. The question of determining the limits of the use of these criteria is also of interest.

The results of this analysis can be applied in scientific research and practical tasks related to CCTV design.

4. Justification of the analysis of scientific research sources

Historically, the first CCTV systems, which appeared in the 1950s, were analog in terms of image transmission and were not subject to industry quality standards. The process of selecting video cameras was largely reduced to a simple comparison of technical characteristics and specifications offered by manufacturers as key selling points. This practice led CCTV developers to make a critical mistake – they did not take into account operational requirements, i.e., the real goals and tasks that the system had to perform.

The level of image detail is one of the key factors when choosing a CCTV camera. This indicator, which reflects the clarity of the object, directly depends on the operational tasks that the system has to perform. It was the need to formalize these tasks that led to the emergence of the concepts of "operational task" and "operational standards." This concept, formulated in the UK in the late 1980s, formed the basis of one of the first global security standards – BS 8418:1987. These standards are of a recommendatory nature.

The BS 8418:1987 standard specified that in the era of analog CCTV, the main criterion for selecting a camera for specific operational needs was the lens. Since the main purpose was to monitor behavior, a test mannequin served as the benchmark for configuring these cameras.

This standard established the first official criteria for CCTV required to perform specific operational tasks. These criteria were based on detail categories determined by the percentage of the frame height occupied by a fulllength image of a person (Table 1).

Table 1
Criteria for solving the operational task of CCTV standard BS 8418:1987

No	Operational task	Operational requirements	Objective	Human height, %
1	Identification	Allows you to see as many details as possible in order to identify an unknown person by their facial features	Gathering evidence, detailed identification of the person	100150
2	Recognition	You can recognize a familiar face, identify clothing features	Tracking acquaintances or confirming their presence at a location	50
3	Detection	Ability to determine with high confidence whether a person is present in the frame	Confirm the presence of a person, determine the direction of movement	10

Due to improvements in analog CCTV resolution, BS 8418:1987 has been replaced by BS 8418:2009. The new edition of the standard has significantly expanded the list of operational tasks to include monitoring, detection, surveillance, recognition, identification, and inspection.

The monitoring task in this classification is basic and involves only assessing the general situation at the facility (e.g., the situation at the entrance or in the lobby). It allows you to identify general changes (e.g., crowds of people) but does not provide the detail needed to distinguish faces and license plate numbers.

The task of detection involves the guaranteed identification of subjects or objects of observation in a controlled area. Observation consists of determining the rough characteristics of the object or subject of observation, followed by monitoring their movements in order to identify unauthorized or potentially dangerous actions.

The goal of recognition is to determine whether a subject or object of observation, or elements thereof, belong to a specific group of objects of observation.

Solving the identification task requires the ability to distinguish fairly small, characteristic details, such as the face of a person entering a facility or the license plate number of a vehicle. Thus, identification is the process of recognizing a subject or object by its inherent or assigned identifying features.

Since, in analog video surveillance systems, resolution is determined by the number of television lines (TVL), the requirements for solving the task set in BS 8418 2009 and 2015 for security systems remained based on what part of the frame vertically should be occupied by the object of surveillance (Table 2).

Table 2
Criteria for solving operational tasks of CCTV standard BS 8418:2015

No	Operational task	Tasks and opportunities	Height of image display in frame relative to total height in PAL cameras, %	Height of image display in frame relative to total height in NTSC cameras, %
1	Monitoring	Monitoring and controlling the crowd	5	5
2	Detection	Guaranteed detection of a person in the frame	10	10
3	Observation	Determining a person’s distinctive features, such as clothing	25	30
4	Recognition	Recognition of people familiar to the operator	50	60
5	Identification	Quality sufficient for human identification	100	120
6	Inspection	100 % identification capability, which eliminates doubt	400	400

In turn, Irish I.S. 199:1987, Australian and New Zealand AS 4806.22006, and other CCTV standards of that time also used similar criteria to solve various operational tasks.

In 2013, the European Committee for Electrotechnical Standardization (CENELEC), seeking to standardize the design of CCTV systems in the European Union, issued standard EN 501327 "Video surveillance systems for security television".

The EN 501327 standard defined the criteria for selecting cameras and lenses, methods for assessing scenes and lighting, as well as requirements for the number and placement of cameras. At the same time, the criteria for solving operational tasks (detection, recognition, identification) fully corresponded to the categories established in the British standard BS 8418:2009.

The main difference between the European standard was the introduction of an alternative key parameter – "pixel density" per unit width of the object being observed. The introduction of this parameter was prompted by the emergence of new types of cameras on the market that functioned differently from traditional analog cameras. The relevance of these pixel densitybased criteria increased with the spread of megapixel digital CCTV systems and remained throughout the transition period when digital systems coexisted with analog ones, until they began to dominate.

With the dominance of IP and HD cameras on the market, the old methods of determining operational requirements based on the percentage ratio of a person's height to the frame have become obsolete. Modern digital systems have moved from measuring resolution in television lines (TVL) to pixels. Requirements are now specified as the number of image pixels per 1 meter of the object at a given viewing distance.

Thus, the criteria that existed in global standards for analog systems were gradually improved until CCTV completely transitioned to new digital systems, where the key metric became the "pixel" rather than the "TV line".

Until recently, the operational tasks assigned to CCTV were described by relevant international standards (Table 3). The requirements for monitoring and detection are the most unified among these standards, while the criteria for performing more complex operational tasks – identification and recognition – vary significantly.

Table 3
Comparison of criteria for operational tasks of CCTV according to global standards

No	Operational task	Minimum required pixel density (pixels/m) accordingly to AS 4806	Minimum required pixel density (pixels/m) according to EN IEC 62676	Minimum required pixel density (pixels/m) according to EN 50132
1	Monitoring	17	12	15
2	Detection	35	30	25
3	Observation	70	60	62
4	Recognition	175	125	125
5	Identification	350	250	250
6	Inspection	1000	1000	1000

Changes in the global political situation, a significant increase in terrorist activity, and the development of artificial intelligence technologies have led to a review of operational tasks and, accordingly, the emergence of IEC/EN 626764: 2024 [1].

The characteristic features of the IEC/EN 626764: 2024 operational standard are presented in Table 4.

The data presented makes it clear that the types of operational tasks have changed and, accordingly, the recommendations for CCTV design.

Table 4
Criteria for solving operational tasks of CCTV according to IEC/EN 626764: 2024

No	Operational task	Tasks and opportunities	Required pixel density (pixels/m)
1	Overview	Detect moving objects and their direction of movement	20
2	Outline	Mark objects and their direction of movement	40
3	Discern	Recognize objects and their movement (distinguish between people, vehicles, and animals)	80
4	Perceive	Identify objects and their movements (no distinction based on gender or personal characteristics)	125
5	Charakcterise	Characterizes the type of person, gait and behavior, as well as the type and category of vehicle	250
6	Validate	Confirm familiar faces, actions can be tracked, vehicle license plates can be recognized	500
7	Scrutinise	Identification of people, recognition of vehicles by model and year of manufacture, license plates are clearly legible	1500

The Overview task should enable the operator to identify an object that has just appeared among other elements of the image on the monitor. Based on the definition of the Overview concept, this task is solved taking into account the actual lighting conditions and is evaluated by the degree of isolation of the control object from the general background of the scene.

The influence of technical parameters of video cameras on the image quality of an object was studied by L. Wastupranata [2], H. Gururaj [3], M. Shumeiko [4], J. Vijaya [5], Vlado Damjanovski [6], John Bigelow [7], and others.

Modeling an image of a group of people [8, 9] using specialized software for the task of "detection" (detection), obtained by cameras with different resolutions, indicated that there was no point in installing a camera with unlimited high resolution for a simple task, since this would not provide additional operational value but would lead to unnecessary costs. In other words, there is a certain limit to the resolution that is sufficient for detecting an object, and exceeding it for this particular task is impractical.

The author notes that there is a limit to the practicality of choosing camera resolution for each size of object on the screen. Thus, using a 4 MP camera does not offer any advantages in terms of "detection" compared to a 2 MP camera.

A study of the impact of camera resolution on subject recognition showed that even at the same focal length (the size of a person on the screen does not change visually), the quality and detail of images vary significantly depending on the resolution. In particular, a camera with a resolution of 2 megapixels provides sufficient image quality for successful detailed "recognition" of a subject, while images from a 4megapixel camera can be used to solve the task of “identification” (establishing identity). Thus, if the subject occupies the same space on the screen, increasing the camera’s resolution increases the level of the operational task that can be solved.

When solving Perceive tasks, you first need to understand the conditions in which the objects of observation are located. If the position of the objects is static they are immobile this is one approach, and to solve such a task, you need to determine the focal length of the lens and the resolution of the camera. If the objects are in motion, the number of parameters required increases. To solve such tasks, you also need to determine the exposure time of the electronic shutter. The exposure time of the electronic shutter must be set in order to obtain a clear, nonblurred image of the moving object in each frame.

The issue of transitioning from traditional image quality measurement in television lines (TVL) to pixels, as well as the concept of changing spatial resolution density, have been highlighted and analyzed in scientific research, in particular in works [10, 6, 7] and [11].

The theory of spatial resolution density (or pixel density) variation is a key tool for designing modern CCTV systems. Its essence lies in determining the distance from the camera to the object at which the number of pixels per 1 meter of the width of the monitored area (linear field of view) reaches a predetermined value.

When solving any operational tasks, it is necessary that the resolution of the monitor screen is not worse than the resolution of the video camera. Changes in spatial resolution values obtained experimentally for cameras with HD image quality on the monitor are illustrated in Figure 1.

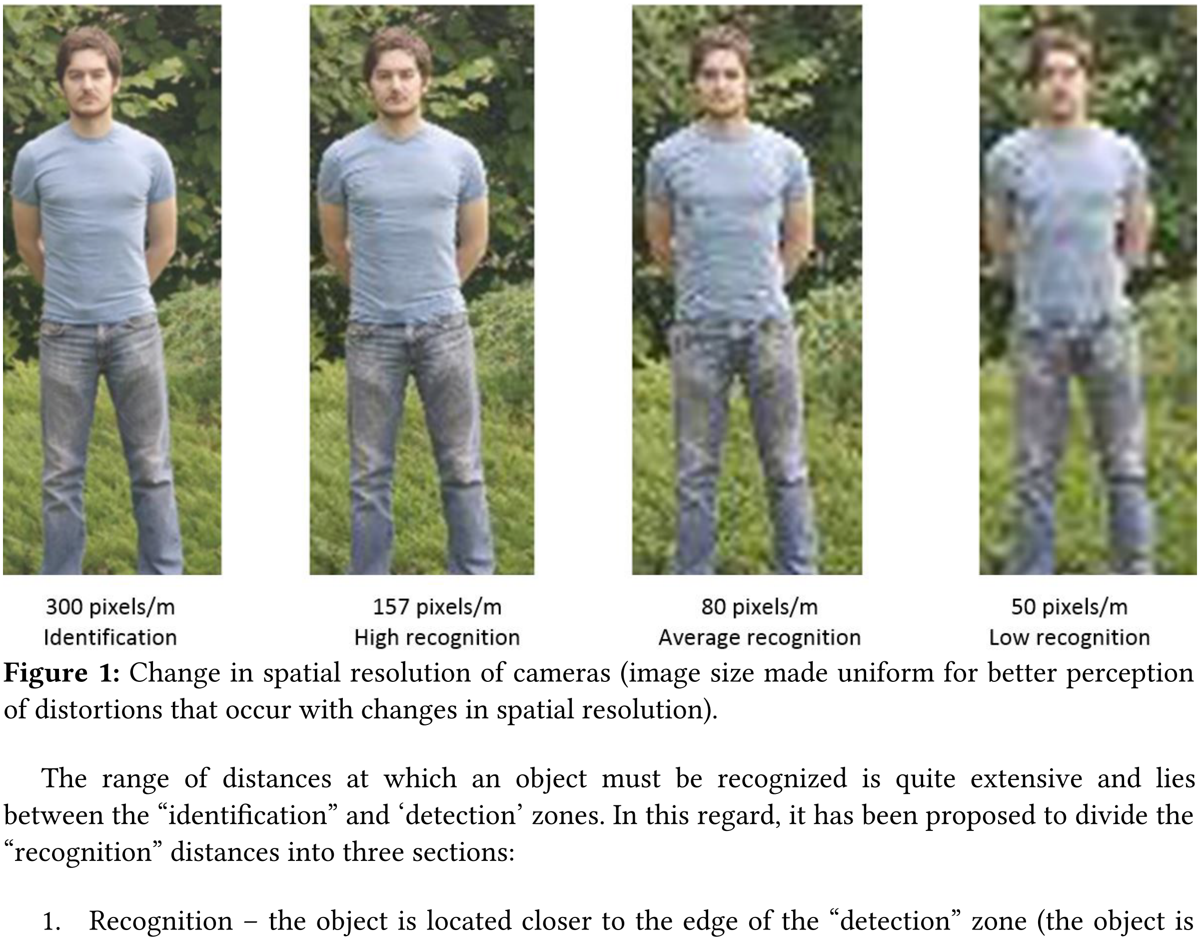

Figure 1: Change in spatial resolution of cameras (image size made uniform for better perception of distortions that occur with changes in spatial resolution).

The range of distances at which an object must be recognized is quite extensive and lies between the “identification” and ‘detection’ zones. In this regard, it has been proposed to divide the “recognition” distances into three sections:

1. Recognition – the object is located closer to the edge of the “detection” zone (the object is small on the monitor and has poor detail).
2. Medium recognition – the size and quality of the object’s display allows its main details to be described.
3. High recognition – the object is located closer to the “identification” area (allows for a clear description of the object).

Thus, analysis of information sources has shown that spatial pixel density is used as a modern criterion for solving various types of operational CCTV tasks, and their values vary in existing standards. The “pixel density” parameter of the camera takes into account the characteristics of the matrix, the lens, and the distance to the object of observation.

It has been established that increasing the resolution of video cameras leads to an increase in pixel concentration per 1 meter of the linear field of the observation scene, which allows for better quality display of observation objects at a significantly greater distance from the camera.

Based on this, it can be assumed that further development of technologies and requirements for CCTV systems may lead to the emergence of matrices with such an extremely high number of pixels that it will change the current classification of CCTV operational tasks. As a result, another question will arise about the further use of existing recommendations for evaluating video surveillance systems.

An analysis of existing studies [3, 12, 13] on the impact of technical parameters of video cameras on the image quality of an object when solving various types of operational CCTV tasks has shown that the issue of excessive design needs to be addressed.

5. Method of implementing

To solve the problem, it was decided to use the theory of spatial resolution change as a basis, which was implemented analytically and through computer modeling in specialized software IP Video System Design Tool 2025.

If the linear field of view Hpv and the number of pixels in the matrix nw across its width are known, then the number of pixels npw per unit width of this field of view can be determined as:

$$n_{n3} = \frac{n_w}{H_{pv}} = \frac{n_w}{L\frac{h_m}{f}}, \tag{1}$$

where L – is the distance from the camera to the object being observed, m; h_m – is the width of the video camera matrix, mm; f – is the focal length of the lens, mm.

To determine the number of pixels across the width of the matrix, you need to take into account its aspect ratio. If N – is the total number of pixels in the matrix and the aspect ratio is 4:3 (width to height), then the number of pixels across the width is determined as:

$$n_w = \sqrt{\frac{n_m}{0.75}}. \tag{2}$$

Hence, the change in spatial resolution can be defined as:

$$n_{sv} = \frac{\sqrt{\frac{n_m}{0.75}}}{L\frac{h_m}{f}}. \tag{3}$$

The focal length, depending on the selected type of task and the distance to the object of observation, is determined as:

$$f = \frac{L h_m n_{sv}}{H_{pv}}. \tag{4}$$

All studies were based on modeling changes in spatial resolution for a video camera with a 12 mm lens and a 1/2inch sensor. The following range of parameters was selected for analysis: the resolution of the sensor was 5 megapixels and was increased in increments of 5 megapixels, while the range of distances to the object was from 5 to 100 meters in increments of 5 meters.

The results of the calculations of spatial resolution changes are summarized in Table 5.

Table 5
Spatial pixel density for a 4:3 format matrix measuring 1/2", f = 12 mm

Matrix resolution n_m, Mpx	5	10	15	20	25	30
Number of pixels n_w, *pcs*	2581	3651	4472	5163	5773	6324
Distance to the object L, m						
5	968	1369	1677	1936	2165	2371
10	484	684	838	968	1082	1185
15	322	456	559	645	721	790
20	242	342	419	484	541	592
25	193	273	335	387	433	474
30	161	228	279	322	360	395
35	138	195	239	276	309	338
40	121	171	209	242	270	296
45	107	152	186	215	240	263
50	96	136	167	193	216	237
55	88	124	152	176	196	215
60	80	114	139	161	180	197
65	74	105	129	148	166	182
70	69	97	119	138	154	169
75	64	91	111	129	144	158
80	60	85	104	121	135	148
85	56	80	98	113	127	139
90	53	76	93	107	120	131
95	50	72	88	101	113	124
100	48	68	83	96	108	118

Modeling of changes in spatial pixel density in IP Video System Design Tool 2025, in order to confirm the obtained calculation data, was carried out using the following data: focal length of the video camera lens 12 mm; distance from the video camera to the human figure model from 20 to 60 m with a step of 5 m; video camera matrix form factor 1/2"; matrix format 4:3; video camera resolution 5Mpx, 10Mpx, 19Mpx. Since there is no 20 Mpx video camera in the program database, a video camera with a resolution of 19 Mpx was used as the closest to the required value.

The human figure model was placed at a distance of 20 m from the video surveillance camera and, after recording the spatial pixel density readings, was moved 5 m further. The spatial pixel density values were obtained together with the formed image of the human model. The simulation results in the selected range are presented in Figure 2.

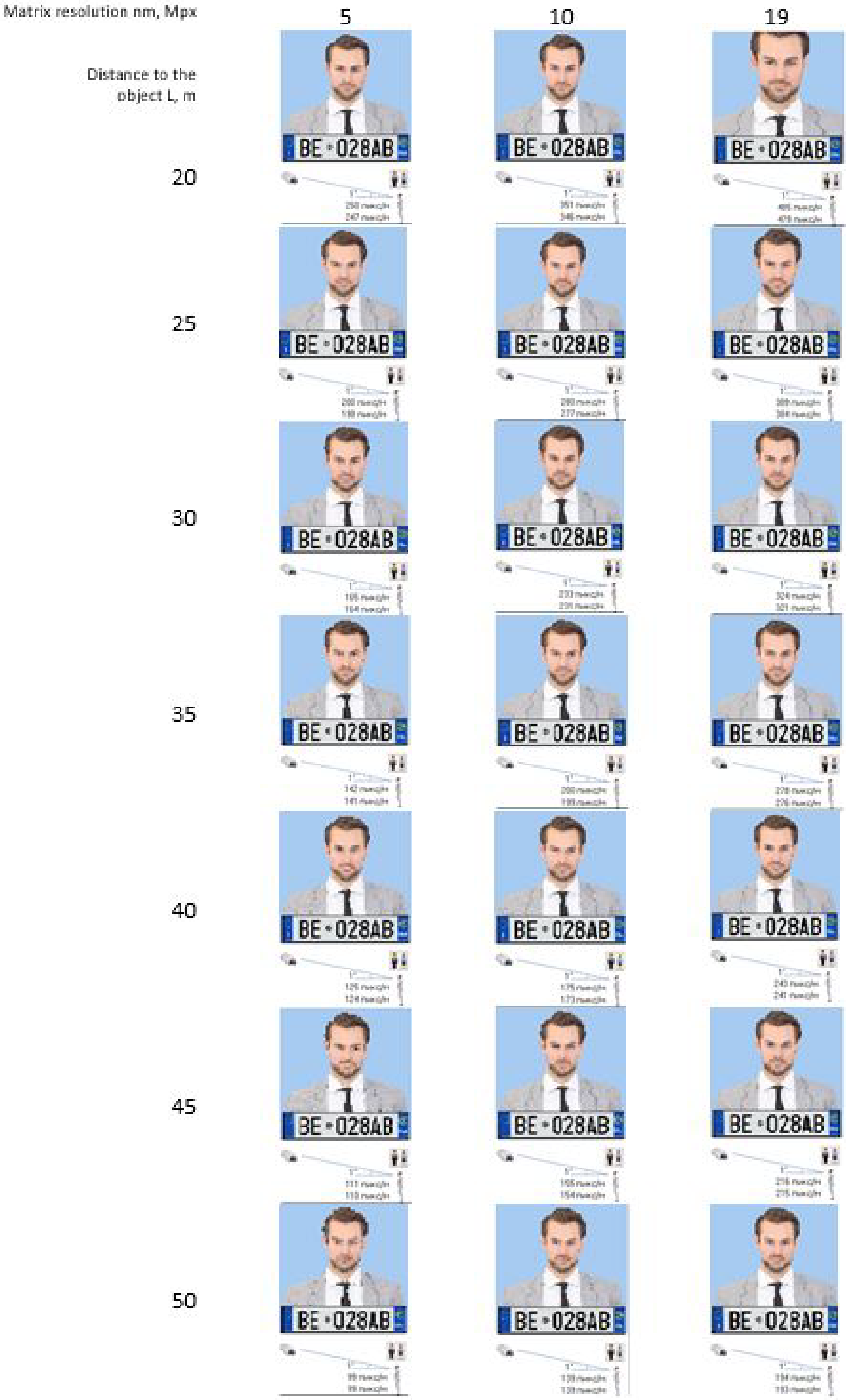

Figure 2: Pixel density simulation results in IP Video System Design Tool 2025.

The corresponding graphical dependencies are presented in Figure 3.

Having determined the type of operational task to be solved for a specific video surveillance area, the designer can use graphical dependencies (Fig. 3) to select the required video camera resolution. This will help avoid excessive design.

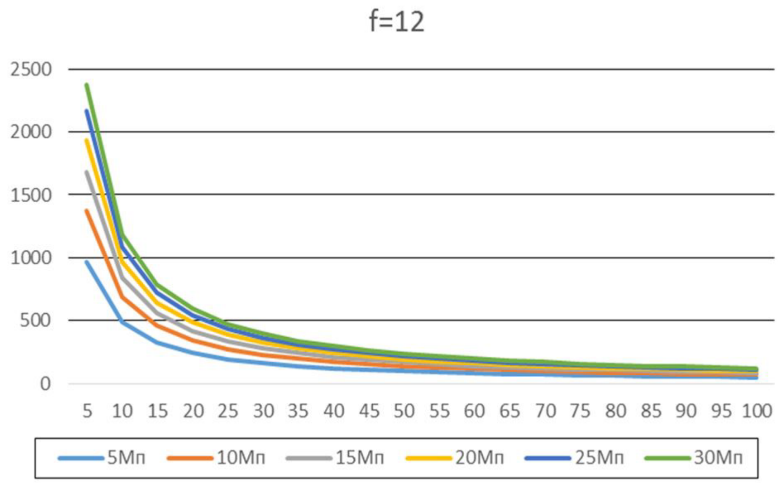

Figure 3: Change in spatial resolution.

6. Conclusions

Comparative analysis of data obtained by calculation and modeling revealed a discrepancy in values at a certain distance from the camera from 5 to 2 pix/m. In general, this indicates that the total deviation of the obtained results is no more than 3.

A person, given his physiological capabilities, can recognize a familiar person at a limited distance. This distance is mainly 3035 m. Thus, the spatial density values recommended in the existing CCTV standards for a certain type of operational task can be fully provided at different distances by video cameras available on the market. The presented data of the research results will help designers to reasonably approach the choice of video cameras, which will reduce the risk of overdesign. The use of a 25 Mpx video camera overlaps the capabilities of the human visual apparatus. Therefore, with the free presence of 25 Mpx cameras on the market, it will be possible to move away from the use of the currently applicable criteria for operational tasks.

It is advisable to rely on the recommended criteria for operational tasks of CCTV standard IEC/EN 626764: 2024 when designing video surveillance systems with integrated video analytics and artificial intelligence technologies that perform a key preventive function. These systems require high scene resolution for analysis and are capable of detecting abnormal behavior in real time that may indicate the preparation of a terrorist act. This early detection capability allows law enforcement agencies to gain situational awareness and respond quickly to potential threats before they escalate into a real incident. Thus, in critical sectors, it will be possible to control their perimeter and perform video analytics of intrusions into the area, monitor safety by checking for the absence of helmets, protective eyewear, etc. At transport hubs, it will be possible to ensure passenger safety, control flows, and counter terrorism by detecting abandoned objects and aggressive behavior, analyzing crowds, and detecting queues. The implementation of these criteria in smart city monitoring systems will improve quality of life, public safety, and the efficiency of city services. This applies to the detection of unauthorized dumping of garbage or the start of a fire, improper parking, the detection of outbreaks of panic or mass riots, etc.

Acknowledgements

Dear organizers of the International Scientific Workshop “Applied Information Technologies and Artificial Intelligence Systems”!

We sincerely thank the organizers for inviting us to participate in the 1st International Scientific Workshop “Applied Information Technologies and Artificial Intelligence Systems.”

We are honored to receive such an invitation. We have reviewed the topics and program of the Workshop with great interest and consider it extremely relevant and important for the development of applied information technologies and artificial intelligence systems.

We are confident that the Workshop will be an excellent platform for exchanging knowledge, experience, and establishing new scientific contacts.

Declaration on Generative AI

During the preparation of this work, the authors used Gemini and Grammarly in order to: Grammar and spelling check and as a smart Search Engine to find related works based on context of conversation. After using these tools/services, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] Video surveillance systems for use in security applications – part 4: application guidelines, 2025. URL: https://webstore.iec.ch/en/publication/83425.

[2] L. Wastupranata, S. Kong, L. Wang, Deep learning for abnormal human behavior detection in surveillance videos – a survey, Electronics 13 (2024). doi:10.3390/electronics13132579.

[3] H. L. Gururaj, I. B. C. S. Senior Member, S. Priya, J. Shreyas, F. Flammini, A comprehensive review of face recognition techniques, trends and challenges, IEEE Access 12 (2024) 107903–107926. doi:10.1109/ACCESS.2024.3424933.

[4] M. Shumeiko, Identification, recognition and detection of people according to the European standard EN 501327, Security Systems, 2015. URL: http://library.tsu.tula.ru/files/elect_periodical/system_securiti3.pdf.

[5] J. Vijaya, S. Chandrakar, P. Shrivastava, A comprehensive review concerning the involvement of artificial intelligence techniques in face recognition system, in: S. Tikadar, J. J. Puga, C. Rodrigues, R. Shaw, A. Hassan, B. K. Patra (Eds.), Practical Applications of Smart Human-Computer Interaction, IGI Global, Hershey, PA, 2025, pp. 69–110. doi:10.4018/9798337363851.ch003.

[6] V. Damjanovski, Case study: evaluating face identification of IP cameras with the Vidi Labs test chart v.5.2, 2014.

[7] J. Bigelow, About pixel densities and what they mean, 2017. URL: https://www.securitysolutionsmedia.com/2017/08/01/aboutpixeldensitiesandwhattheymean/.

[8] Y. MyagmarOchir, W. Kim, A survey of video surveillance systems in smart city, Electronics 12 (2023) 3567. doi:10.3390/electronics12173567.

[9] J. Indhumathi, M. Balasubramanian, B. Balasaigayathri, Realtime videobased human suspicious activity recognition with transfer learning for deep learning, Int. J. Image, Graphics Signal Process. 15 (2023) 47–62. doi:10.5815/ijigsp.2023.01.05.

[10] V. Tsakanikas, T. Dagiuklas, Video surveillance systems – current status and future trends, Comput. Electr. Eng. 70 (2017). doi:10.1016/j.compeleceng.2017.11.011.

[11] D. Uhryn, A. Karachevtsev, T. Terletskyi, O. Kaidyk, M. Talakh, V. Ilin, V. Bogachuk, O. Kaduk, Z. Suranchiyeva, et al., Modern programming technologies in the tasks of identification and classification of military aircraft using machine learning algorithms, in: Proceedings of the Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2024, Proceedings of SPIE, Bellingham, USA, 2024. doi:10.1117/12.3054877.

[12] B. S. Reddy, R. Jha, V. Naware, A. Vattem, A. Hussain, Design and implementation of hybrid lowpower wide area network architecture for IoT applications, J. Ambient Intell. Smart Environ. 16 (2023) 1–13. doi:10.3233/AIS230146.

[13] A. Muhammed, I. Medvedev, N. Gonçalves, VoidFace: a privacypreserving multinetwork face recognition with enhanced security, arXiv preprint arXiv:2508.07960 (2025). doi:10.48550/arXiv.2508.07960.