Int. J. Appl. Math. Comput. Sci., 2022, Vol. 32, No. 1, 147–160
DOI: 10.34768/amcs-2022-0012

TEMPLATE CHART DETECTION FOR STOMA TELEDIAGNOSIS

MARIUSZ SZWOCH[a,*], RAFAŁ ZAWIŚLAK[b], GRZEGORZ GRANOSIK[b], JOANNA MIK-WOJTCZAK[c,e], MICHAŁ MIK[d,e]

[a] Faculty of Electronics, Telecommunications and Informatics
Gdańsk University of Technology
ul. Narutowicza 11/12, 80-233 Gdańsk, Poland
e-mail: szwoch@eti.pg.edu.pl

[b] Faculty of Electrical, Electronic, Computer and Control Engineering
Lodz University of Technology
ul. Stefanowskiego 18, 90-537 Lodz, Poland
e-mail: {rafal.zawislak,grzegorz.granosik}@p.lodz.pl

[c] Faculty of Organization and Management
Lodz University of Technology
ul. Piotrkowska 266, 90-924 Lodz, Poland

[d] Department of General and Colorectal Surgery
Medical University of Lodz
ul. Zeromskiego 113, 90-549 Lodz, Poland
e-mail: michal.mik@umed.lodz.pl

[e] PHIN Consulting Ltd.
ul. Czestochowska 63, 93-115 Lodz, Poland
e-mail: joanna.mik@phin.pl

The paper presents the concept of using color template charts for the needs of telemedicine, particularly telediagnosis of the stoma. Although the concept is not new, the current popularity and level of development of digital cameras, especially those embedded in smartphones, allow common and reliable remote advice on various medical problems, which can be very important in the case of limitations in a physical contact with a doctor. The article focuses on the initial stages of photo processing for the needs of telemedicine, i.e., on the assumptions and the process of designing the appropriate template and detecting it in photos for stoma telediagnosis. Research on the developed algorithms for the location of fiducial markers and reference color fields, carried out on the basis of over 2,000 photos, showed a very high tolerance to scene exposure, lighting conditions and the camera used. The obtained results allowed the initial image intensity normalization of the stoma area as well as correct localization and measurement of changes detected on the skin and the mucosa, which, in the opinion of doctors, significantly increased the diagnostic value of the photographs.

Keywords: stoma, marker detection, telediagnosis, distortion correction.

1. Introduction

Patients with an intestinal stoma often meet substantial psychological and social problems after discharge. Many studies have shown that ostomates are much more depressed, function poorly in the society, and have more disturbed body image than patients who have undergone surgery without a stoma. Stoma negatively affects their interpersonal relationships, health-related quality of life and self-esteem (Boraii, 2017; Hoon *et al.*, 2013; Bulkley *et al.*, 2018). The stoma care systems do not always fully protect these patients; thus, they cannot feel confident in their completely new health situation.

The early postoperative period after discharge is a specific challenge for ostomates as well as the moment

*Corresponding author

when stoma complication arises and the patient has no contact with the stoma nurse or surgeon. We have been struggling with the COVID-19 pandemic for a year and now we know that a direct contact of the patient with medical staff is not always possible.

Fortunately, new technologies allow for the use of tools that enable tele-monitoring for stoma patients. In many cases, photos of the stoma taken by the patient along with a description of the symptoms will allow the correct medical decision to be made during a remote consultation. One of the key factors in this case is the high quality of the photo, in particular, faithful color reproduction and the ability to easily measure selected image fragments. One of the most effective approaches is to use a colored template or card in such situations. The aim of this work is to design such a template for the needs of stoma telediagnosis and to develop methods for its detection in the image.

The research focuses on the initial stages of photo processing, which are: detection of a template chart in a photo based on the selected class of fiducial markers, and then using them to locate reference fields with skin color and texture patterns, as well as straightening the photo by removing any perspective distortion and affine transformations. In order to validate the correctness of the implementation of these tasks, a qualitative and quantitative assessment of the location accuracy was carried out, and selected reference fields were used to improve the image contrast.

The paper is organized as follows. The problems of stoma telediagnosis and fiducial markers detection is detailed in Section 2. In Section 3, the design process of a dedicated template chart is presented. The next section presents the results of experimental verification of selected methods of stencil detection in stoma photos and an image improvement. Finally, some conclusions are given.

2. Background

2.1. Stoma telediagnosis. Representatives of some scientific societies dealing with stoma care emphasize the importance of the development of telemedicine at present. Some reports indicate that in ostomy patients, the use of telemedicine reduces the rate of readmissions and reduces the journeys to the doctor of more than 8 hours on average for a month (Augestad *et al*., 2020). However, telemedicine systems should be constructed to allow the patient to be carried by the same medical team (stoma nurse and surgeon) and not decreasing the chance of contact with a competent physician (Augestad *et al*., 2020; Dinuzzi *et al*., 2021).

The researchers emphasize that postoperative follow-up for ostomates using a mobile application improves the level of psychosocial adaptation and self-esteem compared to patients covered by standard postoperative control. This form of care also reduces the rate of stoma complications (Wang *et al*., 2018; Pata *et al*., 2020).

Recently, research has been conducted on the diagnostic effectiveness of photographs of postoperative wounds to monitor their healing after the patients' discharge (de Heide *et al*., 2017). Each stoma complication has its own characteristic features that can be highlighted in the photo but some stoma patients may have more than one stoma complication simultaneously, that makes the diagnosis based on the photo more difficult. To obtain the photo of the wound with the highest usefulness it should meet three quality criteria:

- *clarity*—the image must be of significant resolution, free from glare and smudges;
- *context*—the wound must be visible in relation to its surroundings skin;
- *completeness*—the entire wound should be photographed without any hidden elements.

Good quality in the above three criteria will allow us to get a picture with proper diagnostic power (Li *et al*., 2019).

The diagnostic usefulness of the photograph was checked, among others-on traumatic wounds and difficult to-heal postoperative wounds, as well as in many other applications. Regardless of the type of medical imaging and the specific application, the use of appropriate methods of pre-processing, segmentation and recognition of images is extremely important (Ciecierski, 2020). A procedure of taking pictures of the stoma was also developed, but it did not validate the quality of the pictures (Buckley *et al*., 2005). In telemedicine, in order to provide proper stoma care, the patient takes a photo of the stoma himself or herself and sends it to the nurse or doctor. This requires working out a detailed procedure for the patient. The regular stoma is not a wound, even though it is not covered with healthy skin but with mucosa. In complicated stomas wounds may appear around the mucosa and only the external areas are covered with normal skin. This makes the variety of periostomal wound images enormous, so we should get the best possible quality of photos, taking into account not only colors of mucosa, but also the shades of gray and shades of the colors of surrounding skin.

The location of peristomal hard-to-heal wounds is an additional important element in the descriptions of a complicated intestinal stoma. It is related to different therapeutic procedures. In telemedicine we have the lack of direct medical examination of the patient and the knowledge of the spatial orientation of complications on the skin around the stoma (the location in individual skin quadrants, location closer to the mucocutaneous fusion),

allows the nurse to offer the patient the best solution, i.e., appropriate dresses or stoma equipment.

To sum up, in remote ostomy diagnostics it is extremely important to reproduce the original colors as well as to maintain the proportions and dimensions of the examined body area. In the case of colors, it is also very important to be able to compare them with a certain set of reference colors as well as a set of textures showing different colors of the skin and the stoma itself. In the case of self-taking photos in non-optimal conditions, it is also important to be able to improve the quality of photos, i.e., improve contrast, sharpness, noise reduction, etc., while maintaining the original colors.

2.2. Stoma-Alert-Final project. The main goal of the Stoma-Alert-Final project is to provide stoma patients and doctors with a reliable and easy-to-use environment for telediagnosis when a visit to the doctor's office is impossible or unnecessary. One of the important elements of this process is the assessment of the color, size and shape of the stoma and the surrounding abdominal wall. As demonstrated by previous studies (Zawiślak *et al*., 2020), the perception of colors among doctors is highly subjective even in direct contact with the patient. The reproduction of the real colors of objects in a photo and the perception of such a photo in a digital system depends on many factors. The most important of them are:

- *lighting quality*, including brightness level, type of light source and its color temperature and uniformity of illumination;
- *camera quality*, including sensitivity, resolution, dynamics and the level of the noise of the photosensitive matrix, brightness and the level of geometric distortion of the optical system and others;
- *camera exposure*, in particular the correctness of white balance, color temperature, shutter speed, aperture, ISO sensitivity level and others;
- *photo composition*, including all the diagnostically important elements and not introducing additional disturbances, e.g., shadows;
- *screen quality*, including the range and accuracy of color reproduction (calibration), brightness settings, resolution, the presence of an anti-reflective layer and others;
- *differences in color perception*, in particular partial or complete color blindness.

In general, it can be assumed that the digital representation of the colors in the photo differs to some extent from the actual colors, and also that the image displayed on the screen does not exactly represent the digital form. In other words, different cameras may record the same scene differently and miscellaneous devices may display the same photo differently. What is more, the same camera may record the same image in color differently in subsequent photos (e.g., with a slightly shifted exposure), and the same device may display the same image differently with various settings (e.g., by changing the settings of brightness, contrast, color temperature, gamma correction, color balance, etc.).

Unfortunately, there is no one-size-fits-all solution to this problem due to its complexity and the number of different factors determining human color perception. The most important and most frequently used solutions include:

- *calibration* of devices allowing the most faithful representation of colors from a given color model on a given device; special color templates and a colorimeter are used for calibration;
- the use of *color charts* for a relative comparison of colors in the same photo, regardless of the device on which they are presented.

2.3. Fiducial marker localization. Detecting 3D objects in an image, and the equivalent estimation of a camera pose, are a fundamental problem in numerous computer vision applications, such as simultaneous localization and mapping (SLAM) in autonomous navigation (Lim and Lee, 2009), augmented/mixed reality (AR/MR) systems (Li *et al*., 2017), image-guided radiotherapy (Nolan and Forde, 2016), computer-aided implant surgery, and many others.

In general, there are two main approaches to solving this problem. In the first, based on the way people perceive the world, research focuses on detecting and recognizing natural features or *landmarks*. Its advantage is the ability to operate in an unchanged environment, but at the cost of a much higher complexity of computer vision and artificial intelligence algorithms (Todt and Torras, 2000). An alternative, commonly used approach is based on the detection of artificial *fiducial markers* placed in the environment. Although some interference with the environment or the object being analyzed is necessary, this approach is generally more accurate, efficient and smart (Muñoz-Salinas *et al*., 2019).

In addition, to precisely determine the location and orientation of objects on the scene, reference markers can also be used to measure them, as well as to encode specific information, e.g., labeling. In this case, a certain binary code is used, as in the case of bar codes or QR codes. A special case here is the detection of markers from a specific set, or *dictionary*, often many in one image. In such a situation, it is important to properly design (encode) the markers in order to maximize the

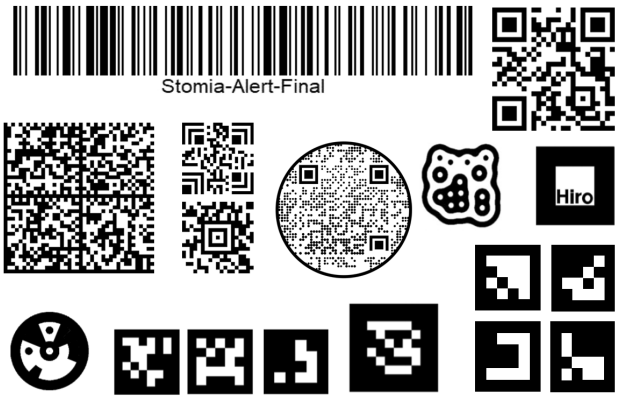


Fig. 1. Example of different fiducial marker types (Linear, QR, Data Matrix, Han Xin Code, Aztec, DotCode, ReacTIVison, ARToolKit, Intersense, AprilTags, ArUco).

inter-marker distance. Moreover, when used as color chart markers, it is also possible to use them for color calibration or to obtain reference colors.

Fiducial markers are used in many different areas of science, technology, and everyday life. In medical imaging, they can be used, among others, to correlate images from two different systems (e.g., PET and MRI), to mark an image, or to actively stabilize the field of view and focus. In the broadly understood metrology, they allow the measurement of size and density. In turn, in photogrammetry, they can be used to calibrate the optical system and determine its parameters. Markers are also used in many AR/MR systems to place virtual objects in real space, e.g., in hybrid games. Robotics, or more generally autonomous vehicles, is an equally important field of application, in which markers are used to locate objects (Sani and Karimian, 2017) or to simplify the SLAM problem (Hoube *et al*., 2016).

Although there are basically no major restrictions on the shape and appearance of fiducial markers, in practice they should be designed for the ease and certainty of their detection in the widest possible range of camera pose and image quality, in particular its resolution, sharpness, contrast, lighting or noise level. For this reason, most fiducial markers contain black and white elements with strong edges and corners.

Many types of markers have been proposed in the literature, some of which have been used in various applications for many years (Fig. 1). The simplest of them are fiducial points for recognizing single correspondence or reference points, e.g., joints in motion capture systems or pose and tracking systems (Tändl *et al*., 2009). Such markers are usually made using light-emitting materials (e.g., LED), reflective or in a specific color, which allows their easy detection in the image. Unfortunately, the detection of their poses is most often based on their mutual spatial relations, which limits their use and can be a relatively complex process. Moreover, markers of this type cannot encode any information.

Another group of markers includes more or less complex graphic shapes that allow for unambiguous identification of their poses based on techniques of detecting geometric primitives or blobs. One can distinguish here in particular circular, d-touch, finger or amoeba markers (Kaltenbrunner and Bencina, 2007).

The next group includes tags composed of prominent key objects or landmarks with a certain topology. These objects can be rings or circles, squares, crosses, lines, and more. The most frequently chosen topology is a regular rectangular grid, which provides a constant and known metric base which also enables camera calibration and reduction of geometric distortions (Kowalski *et al*., 2014). A special case of such a marker of this type is the checkerboard, which, due to the easily and reliably detected corners forming a uniform reference grid, is often used to calibrate devices and determine geometric distortions (Rufli *et al*., 2008). In general, it can be assumed that detecting inner corners as mutually touching opposite squares of a checkerboard is more accurate and stable than detecting individual outside corners. A significant disadvantage of many markers from this group is the limited ability to encode information and some difficulties in creating dictionaries with well distinguished elements.

Very popular markers commonly used to encode product information are bar codes (Hansen *et al*., 2017). Although they can also be used to locate objects, the precise positioning of their poses requires a precise location of the end of the lines, which may be difficult with larger geometrical distortions and lower quality images.

Square-based markers allow easy localization of three or four correspondence points (their corners), as well as for coding information using a combination of white and black internal squares or any graphics. For this reason, this type of marker is now widely used in many applications requiring the precise positioning of a marker (e.g., AR) or camera pose (e.g., SLAM). Initially, square markers allowed encoding information on the basis of any graphics placed in them, e.g., in the popular ARToolKit library (Kato *et al*., 2000). Unfortunately, the main problem of this approach is the use of the template matching technique, which often results in a high false positive rate (Fiala, 2005; Ciążyński and Fabijańska, 2015).

More advanced markers use binary coding of information based on a combination of black and white squares, most often also introducing the ability to detect and correct recognition errors based on redundant bits. This type of markers includes, among others, ARToolKit,

Plus, ARTag, Matrix, AprilTags and ArUco systems (Garrido-Jurado *et al.*, 2016).

As with most tags, one of the main problems is ensuring image rotation resistance while still being able to distinguish all markers in the dictionary. One of the possible solutions is the use of additional markings allowing the recognition of the current marker orientation, as is the case, e.g., in QR codes. However, this complicates the process of recognizing them, and the detection of these landmarks is error prone.

Another solution is to search for such a set of markers that will maximize their mutual distances, taking into account all possible rotations. In this case, you can search the space of all possible tags to find the optimal dictionary or use the generating approach. In both cases, the problem is computational complexity and ensuring a sufficiently large distance between the markers. To solve these problems, various marker generation algorithms have been proposed (Garrido-Jurado *et al.*, 2016).

In many applications, a significant problem of the marker recognition system is the problem of their size (scale) in the image, resulting mainly from the assumption of a large range of detection distances. In this case, multi-scale approaches are proposed. Another interesting approach is the use of multiscale markers such as 'fractal' markers (Romero-Ramirez *et al.*, 2019). The last approach also increases the resistance to marker occlusion to some extent.

3. Template chart design

Due to the problems described in the previous sections, the Stoma-Alert-Final project decision was made to use color card templates to assess the condition of the stoma. This approach eliminates most of the problems related to the reliable color representation and their perception by doctors. However, the cost to be incurred in this case is the need to use an appropriate color chart template, which causes some problems when taking photos and requires its exact positioning in the image with the accuracy of individual fields, or chips, containing reference colors. This section presents the main assumptions about the template and the process of its design.

Determining the target form of the template chart is not trivial and requires taking into account all the initial assumptions and the expected high efficiency and speed of its detection in photographs taken with different cameras from different shots in various lighting conditions. Taking this into account, the template should consist of three main elements:

- a set of markers allowing the efficient location of the template chart in the photos;
- a set of reference color fields (chips) allowing color correction of the photo;

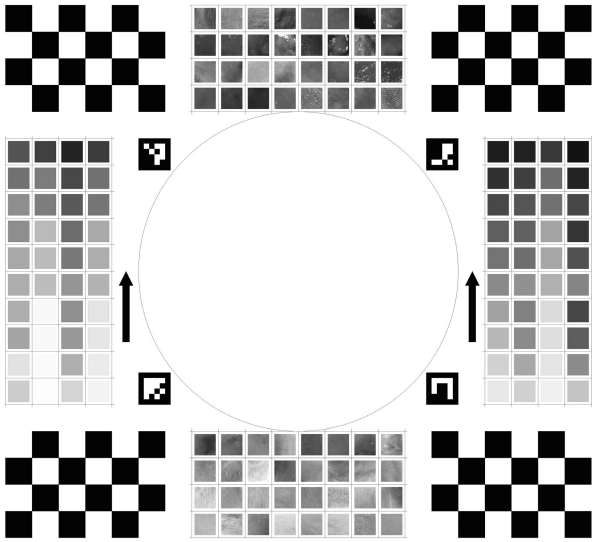


Fig. 2. Proposed color template chart.

- a set of reference fields containing typical skin and stoma textures used by a diagnostician.

Consultations with stoma diagnoses and preliminary studies with patients (Szwoch *et al.*, 2021) showed that for diagnostic purposes an A4 sheet size template with a central hole 12 cm in a diameter with reference skin color and texture palette arranged around its circumference will be sufficient (top and bottom palettes in Fig. 2).

Although the presence of a color template in the photo resolves the main problem of the diagnostician, which is the possibility of comparing the diagnosed patient's skin with the palette of patterns, it does not allow an automatic image quality improvement, in particular white balance correction, light equalization, the removal of geometric distortions or cutting out the central region of interest (ROI).

To obtain such additional possibilities, fields with a reference color palette (including an achromatic palette) should be included in the template (left and right palettes in Fig. 2), as well as additional fiducial markers allowing not only its detection in the image, but also the precise location of all reference fields. An additional advantage in this case will be the ability to determine the pose of the camera at the time of taking the photo, assess its quality, and indicate parts of the skin surface with a specific color.

Choosing the right type of markers requires taking into account the specificity of its application, in particular: lighting conditions, the effectiveness and reliability of marker detection, their number in the image, the need to encode information and many others. Detection of markers in an image can take place at several levels, including salient points or landmarks (e.g., lines, edges, corners, crosses), complex geometric structures (e.g., a

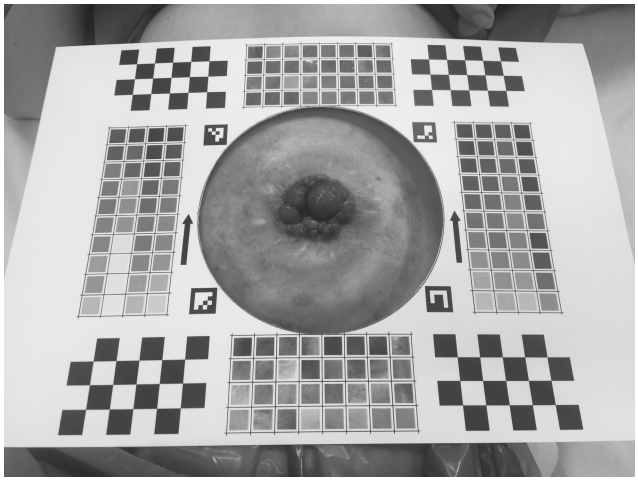


Fig. 3. Example photo from the stoma set.

grid of points or lines, checkerboards), square, circular or other fiducial markers, or any texture patterns (cf. Section 2.3).

In the case of the Stoma-Alert-Final project, flat stiffened templates will be used in a variety of lighting conditions, and the photos will be taken with cameras of different quality. This means the need to use contrasting black and white markers with high detection reliability. There is also no need to encode additional information in the markers; however, square markers are still worth considering for speed and reliability in their detection.

In theory, four markers, e.g., in its corners, should be sufficient to clearly locate a perfectly flat template in a perspective view. Since affine and perspective transformations maintain collinearity and proportionality, it is possible to use simple geometric methods to determine the position of any reference fields based on the detected markers. Unfortunately, in practice, even the smallest bending of the template or defects in the optical path of the camera causes geometric distortions that are difficult to predict, which makes it difficult or impossible to correctly detect the color reference fields. In this case, you can increase the number of markers or use additional positioning elements, e.g., a line grid, contrasting field boundaries, etc.

The situation is very similar when it is necessary to straighten the image by determining and inverting the perspective transformation. Although in theory the homography can be determined on the basis of 4 corresponding points, in practice the mapping determined in this way is very susceptible to even slight errors in determining points in the image. Therefore, many practical applications use a greater number of fiducials.

Taking into account the above aspects, the initial concept of the template assumes the use of three types of markers (Fig. 2):

- four square markers, located on the diagonals of the square described on the circle delineating the central hole;
- four checkerboards, located in the corners of the template;
- a regular grid of horizontal and vertical lines, the intersections of which are marked by the vertices of the colored reference fields which can also be used as a reference measure for calculating the distance in the image.

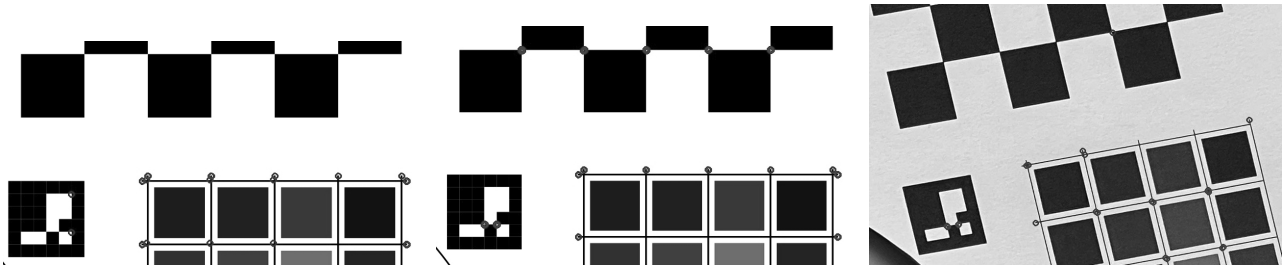
Although such a design is redundant, it will allow the full use of the potential of selected types of markers, especially in the case of unusual framing, difficult lighting conditions, significant bending of the template or the use of a poor-quality camera. Square markers will allow us to quickly locate the template chart in the image and determine its orientation. In turn, the checkerboard will allow a certain determination of the appropriate number of reference points allowing the determination and inversion of the perspective and affine transformations. Finally, the square grid of lines enables easy positioning of individual pattern fields and precise determination of the position of the reference colored fields in the event of geometric distortions. In the case of large distortions, it will also be possible to use line tracking algorithms to determine their actual course. The exact parameters of each type of marker will be determined during the experiments described in the next section.

4. Experiments

4.1. Template charts data set. In order to verify and validate the algorithms for detection of markers and processing of stoma photos with the use of color templates, a collection of over 2000 photos taken by different people, with different cameras, and in various lighting conditions, was created. The photos were taken among both stoma patients in a doctor's office and healthy volunteers (Fig. 3). Due to the iterative development of the template design, all photos are grouped into several subsets based on the same version of it.

The basic assumptions included:

- use of a stiffened flat template in various variants allowing testing and verifying the utility of the templates;
- use of cameras built into the smartphone in basic, automatic settings; not using digital zoom or special smartphone functionalities, such as HDR;
- turning off the flash, unless there is no other option to illuminate the scene;



(a) (b) (c)

Fig. 4. Corner detection using the Harris descriptor (circles indicate corners found): no checkerboards corners with $k = 0.05$ (a), missing grid crosses $k = 0.09$ (b), missing corners on a photo from a data set (c).

- placing the template on the patient's abdomen with the arrows pointing to his/her head;
- good illumination of the template, preferably with diffused or directional (or conical) light in a position that does not cause the appearance of a shadow;
- setting the camera at such a distance that the frame covers the entire template;
- setting the frame in such a way that each checkerboard is placed in a separate quarter of the photo, i.e., the middle section of the photo sides should not run through any checkerboard.

The collection contains over 100 photos of ostomy patients taken by doctors and nurses. Other photos, taken by more than 200 volunteers, include:

- two photos taken by themselves in a lying position with the front (selfie) and rear cameras;
- two photos by themselves in a standing position using the front (selfie) and rear cameras;
- a set of six photos taken by another person, with an increasing angle of deviation of the camera's optical axis from normal to the template in the direction of any (left or right) side, obtaining deviation angles from 0° degrees with optical axis perpendicular to the template (Fig. 3) to approximately 75° (the optical axis near the plane of the template).

As the main problem at this stage of the research was the correct location of the template under various conditions, several different versions of the template chart were used, differing mainly in the dimensions of the checkerboards and with different sets of reference colors.

The above collection is gradually supplemented by actual photos of the stoma taken in the doctor's office by nurses and by the patients themselves. It will act as a validation set at further stages of the project related to color reconstruction and telediagnosis.

In the following sections, we will describe the steps of analyzing stoma photos using a color template chart. The test application was programmed in C/C++ in the Microsoft Visual Studio 2019 environment with the additional use of the OpenCV 4.5 library and compatible extensions.

4.2. Marker detection. The effective location of the base markers allows further image processing, in particular the location of the entire template, the determination of a perspective transformation matrix, the image straightening, the location of color reference fields and skin patterns, as well as the determination of the central area containing the stoma image.

The designed template enables the use of many marker detection methods on different levels, starting from grid lines or crosses, through checkerboards' inner corners, to squared fiducial markers. These methods are not mutually exclusive; on the contrary, they can be used to achieve more accurate results in the case of lower quality images, e.g., heavily noised or distorted.

As part of the experiment, three methods of template location were investigated. One of the basic assumptions was to check the effectiveness of their operation with as little pre-processing of the image as possible. The first approach used the Harris descriptor H to detect corners in an image, particularly checkerboard corners and grid line intersections.

$$H(p,q) = \det M(p,q) - k \cdot (\mathrm{tr} M(p,q))^2 , \qquad (1)$$

where k is an empirically determined constant (usually $k \in [0.04, 0.06]$) and M is a structure tensor (the second moment or the autocorrelation matrix) defined by the partial pixel intensity derivatives I_x and I_y in the x and y directions at point (p,q), the weighting function w (typically an isotropic, circular Gaussian), in a window W of specified size:

$$M = \sum_{(p,q)\in W} w(p,q) \begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix} . \qquad (2)$$

In general, the Harris corner detector offers good repeatability with changing lighting and rotation.

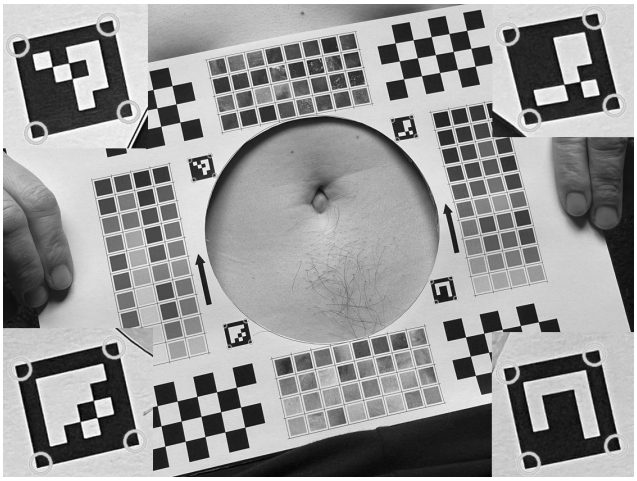


Fig. 5. Example detection of squared fiducial markers using the ArUco library. The enlarged fragments of detected markers indicate some problems with accurate corner detection.

Unfortunately, preliminary testing the Harris descriptor showed its relatively low efficiency for the analyzed images and a high sensitivity to its parameters. As a result, appropriate parameters should be manually selected for different images in order to obtain a clear answer (Fig. 4). In many cases, a minimal change in the k value results in a loss of detected corners or a huge increase in the detected image abnormalities.

Because the Harris detector efficiency was much higher for good quality scanned images than for a photo, it can be assumed that using appropriate photo pre-processing algorithms (e.g., sharpening, denoising, lighting equalization etc.), it would be possible to achieve almost the same high efficiency for certain fixed parameter values. Moreover, by using the even distribution of mesh nodes, it could be tempted to filter out excessively detected corners (e.g., with a spatial comb filter) and fill in the missing ones. Unfortunately, the assumptions of the experiment limited the additional processing, which makes it necessary to discard the global version of this method for the template detection. However, the Harris detector can be used locally in subsequent steps to precisely locate vertices whose rough position has been determined by other methods such as interpolation or extrapolation from detected landmarks.

The second approach focused on detecting square fiducial markers located at the corners of the central region of interest. The experiment used the ArUco library, which is an extension of the OpenCV library and currently offers one of the best algorithms for detecting square markers and four markers from the predefined dictionary `DICT_4X4_250` (Romero-Ramirez *et al.*, 2018). The conducted experiment has shown that it is also very efficient and allows trouble-free real-time recognition of four markers in a wide range of shooting angles (Fig. 5). The exact position of 16 corners of the four markers should allow us to mark all the important points of the template, as well as to correct perspective and affine distortions. Unfortunately, in practice, for some markers there were minor errors in the location of their corners (e.g., the upper right marker in Fig. 5). In combination with even slight disturbances in the pattern planarity, it causes significant errors in the location of mesh nodes (reference fields) and in straightening the image. For this reason, the ArUco library will not be used for these tasks; however, the observed effectiveness of marker detection inclines to use it at the stage of initial location of the template in the photo.

The third approach explored was the detection of the checkerboards present in the image. Qualitative tests were carried out on the entire data set, while quantitative tests were carried out on a subset of 402 photos taken by 67 volunteers using the same template version. The tests carried out showed a 98% of checkerboards detection efficiency for all test photos in which the deviation of the camera's optical axis from the normal template did not exceed 45° (Table 1). Moreover, for many photos it was also effective in the case of a much larger deviation (Table 1). Such a high tolerance of deviation is particularly important in the case of self-taking pictures of the stoma by patients, because in this case it is impossible to obtain a perfect exposure of the image. The checkerboard detection also proved to be extremely resistant to changing lighting conditions, as well as minor problems with image sharpness.

The analysis of the factors causing the failure to detect checkerboards confirmed the knowledge possessed during the creation of assumptions for the acquisition process. They include, among others: highly reflective sheets of paper or paint (Fig. 6(a)) and taking a picture at an acute angle and, related to it, the lack of sharpness due to a shallow depth of field (Fig. 6(b)).

There are also certain additional requirements for the detection of checkerboards, but these are not difficult to meet. The most important of them is the need to capture the entire (not covered or cut) checkerboard in the picture. Under controlled conditions of taking pictures, however, this does not seem to be a problem; on the contrary—it can be a necessary condition for recognizing the picture as correct.

Another problem with using checkerboards is detecting only one checkerboard of a given size in an image. A simple solution is to use four different checkerboards in the template chart. Another solution is to split the image into separate regions (four quarters in this case) and detect one checkerboard in each of them. In the latter solution, the checkerboards may be identical in their dimensions (Fig. 2), but there is an additional

requirement that each checkerboard must be located in a separate analysis region. The practice has shown that this condition is very easy to meet and no such problem appeared in any photo from the test set.

Both approaches were tested in the research. When four different sizes of checkerboards were used, they were detected throughout the image without any problems. Similarly, in the case of dividing the image into quarters, no problems with checkerboard detection were found. Because we want to process the image only once (not four times) and there is no problem with dividing the image into quarters, this solution was selected for further experiments.

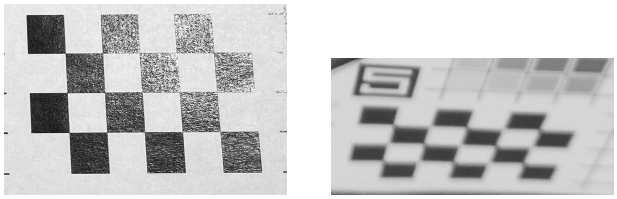
Subsequent studies focused on determining a quasi-optimal size of the checkerboards in the template chart. Since in the next stages of the analysis the checkerboards will be used to determine the position of the reference fields, their larger size means a more precise location of the template and the possibility of defining more such fields. On the other hand, checkerboards that are too large increase the chance of their being accidentally obscured, and also reduce the tolerance of variations when taking a photo. Therefore, as a result of research, a certain compromise size of the checkerboard should be established.

The conducted research for all combinations of sizes ranging from 4×4 to 6×8 fields showed that the size of the checkerboards does not affect the effectiveness of their recognition, and smaller checkerboards are detected a bit faster. Since rectangular checkerboards with even dimensions do not have any axial symmetry, they allow an unambiguous determination of their spatial orientation, which is important when detecting a rotated image. Therefore, a 4×6 checkerboard was considered a quasi-optimal dimension thus determining the appearance of the template for further research (Fig. 2). An example of the effective straightening of the photo is shown in Fig. 7(a).

4.3. Perspective distortion removal. One of the important diagnostic values of an image is the absence of perspective distortion, which allows the measurement and evaluation of the size of significant important dimensions and distances. For this reason, estimating the perspective transformation of a photo and correcting it is one of the key stages of template processing.

The transformation between the planar template sheet π and the image plane on the photo π' can be defined as a planar homography H,

$$s \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad (3)$$



(a) (b)

Fig. 6. Examples of undetectable checkerboards due highly reflective paint (a), out of focus caused by an excessively sharp shooting angle (b).

Table 1. Percentage of photos with the number of checkerboards found as a function of the deviation angle of the camera's optical axis from the normal vector to the template.

	0°	15°	30°	45°	60°	75°
4	100%	100%	97%	94%	97%	39%
3	–	–	3%	3%	–	9%
2	–	–	–	3%	3%	33%
1	–	–	–	–	–	6%
0	–	–	–	–	–	12%

Table 2. ROI detection error [pix] as a function of the deviation angle of the camera's optical axis from the normal vector to the template (calculated in case of 4 checkerboards found).

	0°	15°	30°	45°	60°	75°
Mean	2.72	2.74	2.88	3.58	4.02	5.76
Std. dev.	1.78	1.78	2.16	2.57	2.91	3.87

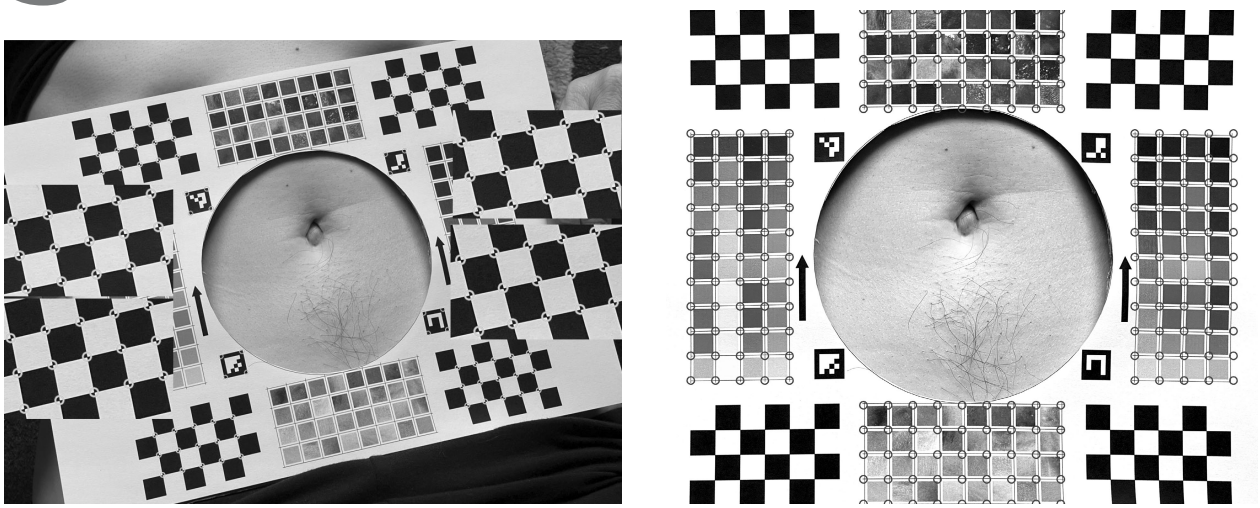
where (x, y) is a point from the template chart plane π, (x', y') is a point from the image plane π', and s is a scale factor.

To enforce 8 degrees of freedom (DOF) for a 9-element homography matrix, it is possible to set $h_{33} = 1$ or impose the unit vector constraint

$$\sum_{i,j=1}^{3} (h_{ij})^2 = 1. \quad (4)$$

Although in theory the homography can be determined on the basis of only 4 pairs of the corresponding points, as mentioned before, in practice the mapping determined in this way is very susceptible even to slight errors in determining points in the image. Therefore, many practical applications use a greater number of fiducials.

Proper detection of at least one checkerboard in the image allows determining the global perspective transformation of the entire image. Of course, detecting all four checkerboards significantly improves the accuracy of estimating this transformation, especially when the planarity of the template has been disturbed. The perspective transformation is described with elements of



(a) (b)

Fig. 7. Examples of checkerboard detection in stoma photography (circles indicate the inner corners of the detected checkerboard) (a), and removal of the perspective transformaction (circles denote grid nodes found) (b).

a 4×4 unified 3D transformation matrix. In order to correctly determine the spatial projection, the coordinates of the corners of the checkerboards in the template and in the analyzed photo should be determined. In addition to maintaining the correct topography and order of individual vertices, the image scale should be correctly selected.

Since the homography is an invertible transformation, the inverse matrix H^{-1} can always be determined and used to remove a perspective distortion to obtain a straightened or dewarped image. An example of the effective straightening of the photo is shown in Fig. (7b).

4.4. Location of grid nodes. The main purpose of the template detection is the precise location of the central ROI containing the stoma image and individual color patterns that allow the original colors of the photo to be recreated. The location of mesh nodes is possible both in the original and straightened image, but in the latter case it will be much easier.

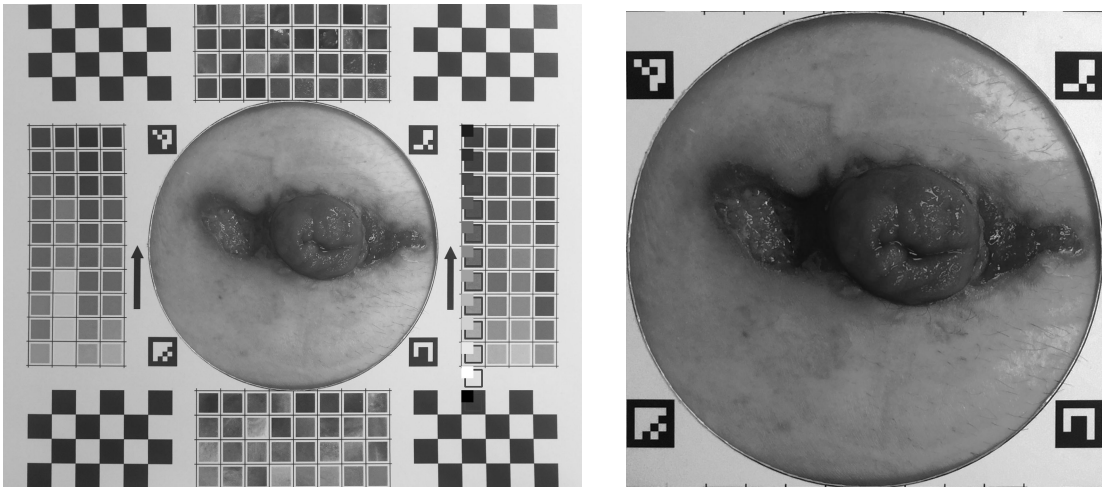
The research carried out on the test set has shown that in the case of a stiffened template, nonlinear distortions in the straightened image are minimal and allow precise determination of reference colored fields. An example of determining interpolated mesh nodes and reference fields is presented in the Fig. 7(b). For larger geometric distortions, local methods for finding mesh nodes as well as curve tracing methods can be used.

In order to estimate the quantitative effectiveness of the grid nodes detection, a measurement of the accuracy of determining the corners of all four reference fields was carried out on the previously used subset 402 of manually labeled photos. The inverse homography matrix H^{-1} was used to automatically determine the position of 16 corners in the input images, and the Euclidean metric was used to measure the distance with the corners indicated by hand. In practice, position deviations of 3-4 pixels allowed a problem-free location of the reference fields, hence 4 pixels were adopted as the upper limit.

The tests carried out on the test set showed 98% efficiency of detecting the central hole and colored reference fields for the stiffened templates within the range of $\pm 45^\circ$ deviations of the camera's optical axis from the standard template (Table 2, Fig. 9). Such high efficiency for highly differentiated images means high stability of the applied approach and confirms its rightness.

The total time of detection of markers, location of reference areas and removal of geometric distortions depends on several factors, including mainly the size of the input images and their quality—for images of low quality, the time of pre-processing and detection of markers increases due to the iterative algorithms used. For the tested set of 402 photos, the average processing time was about 1.5 seconds (in the range of 0.7–3.5 seconds) on a notebook with an Intel Core i7-8850H CPU running at 2.6 GHz and 32 GB RAM.

4.5. Initial contrast correction. In order to assess the effectiveness of the detection of reference colored fields and to initially assess their suitability for improving the quality of stoma images, the initial image of contrast improvement (dynamic range) was performed on the basis of a set of reference achromatic fields. One of the characteristics of low quality cameras (and such unfortunately are built into smartphones) is a low tonal range, i.e., the inability to represent details in the full color



(a) (b)

Fig. 8. Example of center ROI processing the original, low contrast photo (small squares placed on referenced chips indicate the 'ideal', linear intensity levels) (a), the center ROI with a stoma image after contrast correction (b).

and brightness range. It is clearly visible, among others, in Fig. 8(a), where the color representing the white areas is light gray and the black areas is dark gray. The analysis of the actual brightness levels in the original photo shows not only that the contrast is significantly reduced (almost twice), but also that it is significantly nonlinear (Fig. 10).

The first approach uses the simplest linear contrast enhancement function

$$I'(x,y) = C(aI(x,y) + b), \tag{5}$$

where $I(x,y)$ and $I'(x,y)$ are the intensity levels of the original and resulting pixel at (x,y), respectively, a is the contrast coefficient (*gain*), b is intensity bias, and $C(\cdot)$ is the clipping function.

This operation can be performed in the simplest case by setting the brightness as the average of RGB channels, but it may cause a color change in the image, which is unacceptable in telediagnosis of stoma. Therefore, the operation of enhancing the contrast should be performed only on the brightness channel in one of the color models distinguishing such a channel, e.g., HSV.

Based on the brightness of the achromatic fields on the template card, one can approximate the actual brightness line of the photo minimizing the deviation error based on the selected metric, e.g., Euclidean. Assuming that the slope coefficient of the determined line is A, and the lowest brightness level (of the black field) is B, the brightness enhancement function will take the form

$$I'(x,y) = C\Big(\frac{1}{A}I(x,y) - B\Big). \tag{6}$$

An example of the contrast enhancement effect is shown in Fig. 8, as well as in Fig. 10. There is a significant improvement in contrast, especially in the stoma area, and much smaller differences in the brightness of the achromatic fields in the image compared to the 'ideal' level. Obviously, the obtained brightness characteristic is not ideal due to the nonlinear nature of the brightness distortions caused, by the camera sensor, developing a raw photo and other factors. This indicates the need to use more advanced, nonlinear methods of improving the contrast and colors in the image, which will be the subject of future research.

In conclusion, the conducted experiment showed the effectiveness of locating colored reference fields and their usefulness in improving the quality of the original photo.

5. Conclusion and future work

This paper proposes the use of template charts for reliable telediagnosis of the stoma. The proposed template contains: reference color fields that are to enable reliable reconstruction of the colors from the photo on the monitor used by the diagnosing doctor, reference photos of skin fragments, as well as three levels of fiducial markers that are to enable the correct detection of the template in photos of the stoma taken with different cameras in different lighting conditions. The conducted research has shown that when using a stiffened, flat template, four checkerboards placed in the corners of the template are sufficient to carry out the above task. Checkerboards over a wide range of viewing angles and lighting conditions provide an accurate and reliable number of nodal points to remove perspective and affine distortions and precisely locate template reference fields. However, redundant square markers can be used in future works for quick coarse detection of a pattern in an image, and a line grid for the precise location of reference fields in the case of significant geometric distortions.

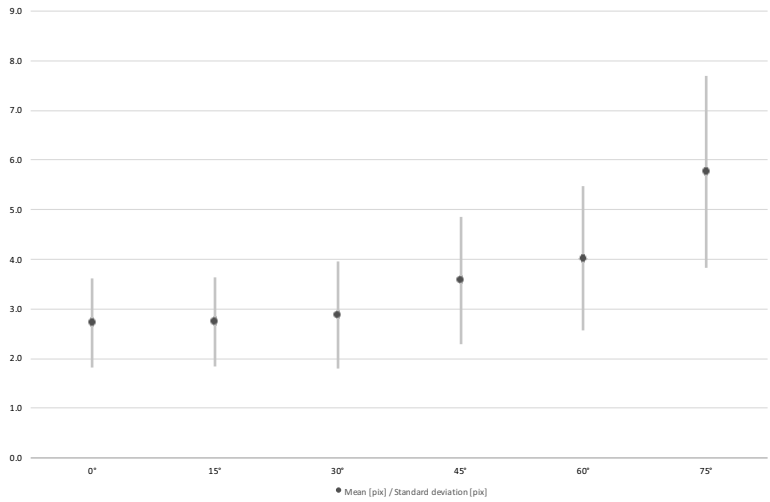


Fig. 9. Average ROI detection error [pix].

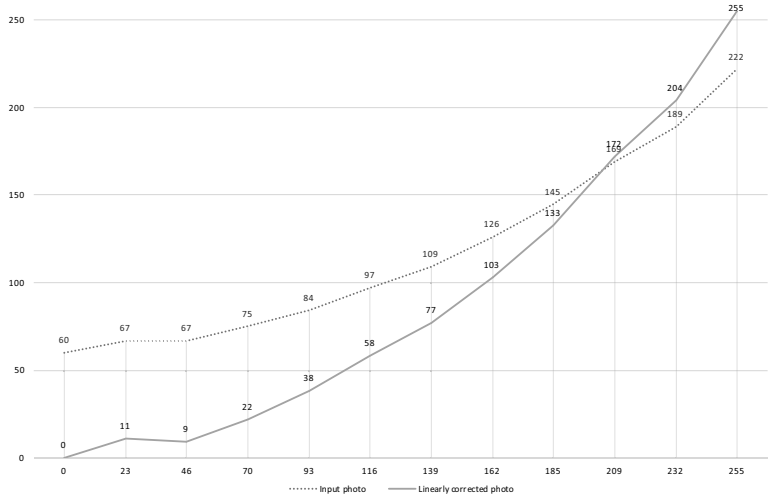


Fig. 10. Brightness transformation function before and after a contrast improvement using a linear function.

The future work will focus on further image processing to improve image quality and, in particular, to reliably represent the stoma area, as well as on the analysis of more flexible templates for which it will not be possible to use homography to remove perspective distortion. In this case, local descriptors will be used to allow accurate detection of mesh points and line tracking algorithms.

Acknowledgment

This work was supported by the National Centre for Research and Development under the grant no. POIR.04.01.04-00-0107/19.

References

Augestad, K., Sneve, A. and Lindsetmo, R. (2020). Telemedicine in postoperative follow-up of stoma patients: A randomized clinical trial (the stompa trial), *British Journal of Surgery* **107**(5): 509–518, DOI: 10.1002/bjs.11491.

Boraii, S. (2017). A descriptive study to assess quality of life in Egyptian patients with a stoma, *Ostomy Wound Manage* **63**(7): 28–33, PMID: 28759426.

Buckley, K., Adelson, L. and Hess, C. (2005). Get the picture! Developing a wound photography competency for home care nurses, *Journal of Wound, Ostomy and Continence Nursing* **32**(3): 171–177, DOI: 10.1097/00152192-200505000-00005.

Bulkley, J., McMullen, C., Grant, M., Wendel, C., Hornbrook, M. and Krouse, R. (2018). Ongoing ostomy self-care challenges of long-term rectal cancer survivors, *Support Care Cancer* **26**(11): 3933–3939, DOI: 10.1007/s00520-018-4268-0.

Ciecierski, K.A. (2020). Mathematical methods of signal analysis applied in medical diagnostic, *International Journal of Applied Mathematics and Computer Science* **30**(3): 449–462, DOI: 10.34768/amcs-2020-0033.

Ciążyński, K.A. and Fabijańska, A. (2015). Detection of QR-codes in digital images based on histogram similarity, *Image Processing and Communications* **20**(2): 41–48.

Dinuzzi, V., Palomba, G., Minischetti, M., Amendola, A., Aprea, P., Luglio, G., De Palma, G. and Aprea, G. (2021). Telemedicine in patients with an ostomy during the COVID-19 pandemic: A retrospective observational study, *Wound Management & Prevention* **67**(1): 12–17.

Fiala, M. (2005). Comparing ARTag and ARToolkit Plus fiducial marker systems, *IEEE International Workshop on Haptic Audio Visual Environments and their Applications, Ottawa, Canada*, pp. 147–152.

Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. and Medina-Carnicer, R. (2016). Generation of fiducial marker dictionaries using mixed integer linear programming, *Pattern Recognition* **51**: 481–491, DOI: 10.1016/j.patcog.2015.09.023.

Hansen, D.K., Nasrollahi, K., Rasmussen, C.B. and Moeslund, T.B. (2017). Real-time barcode detection and classification using deep learning, *International Joint Conference on Computational Intelligence, Funchal, Madeira, Portugal*, pp. 321–327.

de Heide, J., Vroegh, C.J., Szili Torok T., Gobbens, R.J., Zijlstra, F., Takens-Lameijer, M., Lenzen, M.J., Yap, S.C. and Scholte Op Reimer, W.J.M. (2017). A pilot feasibility study of telemedical wound assessment using a mobile phone in cardiology patients, *Journal of Cardiovascular Nursing* **32**(2): E9–E15, DOI: 10.1097/JCN.0000000000000377.

Hoon, L., Chi Sally, C. and Hong-Gu, H. (2013). Effect of psychosocial interventions on outcomes of patients with colorectal cancer: A review of the literature, *European Journal of Oncology Nursing* **17**(6): 883–8913, DOI: 10.1016/j.ejon.2013.05.001.

Hoube, S., Droeschel, D. and Behnke, S. (2016). Joint 3D laser and visual fiducial marker based SLAM for a micro aerial vehicle, *2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), Baden-Baden, Germany*, pp. 609–614.

Kaltenbrunner, M. and Bencina, R. (2007). reacTIVision: A computer-vision framework for table-based tangible interaction, *1st International Conference on Tangible and Embedded Interaction, TEI'07, Baton Rouge, USA*, pp. 69–74, DOI: 10.1145/1226969.1226983.

Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K. and Tachibana, K. (2000). Virtual object manipulation on a table-top AR environment, *IEEE and ACM International Symposium on Augmented Reality (ISAR 2000), Munich, Germany*, pp. 111–119, `http://www.hitl.washington.edu/artoolkit/`.

Kowalski, M., Kaczmarek, P., Kabaciński, R., Matuszczak, M., Tranbowicz, K. and Sobkowiak, R. (2014). A simultaneous localization and tracking method for a worm tracking system, *International Journal of Applied Mathematics and Computer Science* **24**(3): 599–609, DOI: 10.2478/amcs-2014-0043.

Li, M., Howard, D. and King, R. (2019). A picture tells a thousand words smartphone-based secure clinical image transfer improves compliance in open fracture management, *Injury* **50**(7): 1284–1287, DOI: 10.1016/j.injury.2019.05.010.

Li, W., Nee, A. and Ong, S. (2017). A state-of-the-art review of augmented reality in engineering analysis and simulation, *Multimodal Technologies and Interaction* **1**(3): 17, DOI: 10.3390/mti1030017.

Lim, H. and Lee, Y. (2009). Real-time single camera SLAM using fiducial markers, *2009 ICCAS-SICE, Fukuoka, Japan*, pp. 177–182.

Muñoz-Salinas, R., Marín-Jimenez, M. and Medina-Carnicer, R. (2019). SPM-SLAM: Simultaneous localization and mapping with squared planar markers, *Pattern Recognition* **86**: 156–171.

Nolan, C. and Forde, E. (2016). A review of the use of fiducial markers for image-guided bladder radiotherapy, *Acta Oncologica* **55**(5): 533–538, PMID: 26588169.

Pata, F., Bondurri, A., Ferrara, F., Parini, D., Rizzo, G. and MISSTO (2020). Enteral stoma care during the COVID-19 pandemic: Practical advice, *Colorectal Disease* **22**(9): 985–992, DOI: 10.1111/codi.15279.

Romero-Ramirez, F., Muñoz-Salinas, R. and Medina-Carnicer, R. (2018). Speeded up detection of squared fiducial markers, *Image and Vision Computing* **76**: 38–47, DOI: 10.1016/j.imavis.2018.05.004.

Romero-Ramirez, F., Muñoz-Salinas, R. and Medina-Carnicer, R. (2019). Fractal markers: A new approach for long-range marker pose estimation under occlusion, *IEEE Access* **7**: 169908–169919.

Rufli, M., Scaramuzza, D. and Siegwart, R. (2008). Automatic detection of checkerboards on blurred and distorted images, *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, France*, pp. 3121–3126.

Sani, M.F. and Karimian, G. (2017). Automatic navigation and landing of an indoor AR drone quadrotor using ArUco marker and inertial sensors, *2017 International Conference on Computer and Drone Applications (IConDA), Kuching, Malaysia*, pp. 102–107.

Szwoch, M., Zawiślak, R., Mik, M., Mik-Wojtczak, J. and Granosik, G. (2021). Stoma-Alert-Final—Development of IT system model supporting the process of diagnosing, treating and rehabilitating patients with stoma (image part), Project no. POIR. 04.01.04-00-0107/19, *Technical report*, PHIN Consulting and Lodz University of Technology, Lodz.

Tändl, M., Stark, T., Erol, N.E., Löer, F. and Kecskeméthy, A. (2009). An object-oriented approach to simulating human gait motion based on motion tracking, *International Journal of Applied Mathematics and Computer Science* **19**(3): 469–483, DOI: 10.2478/v10006-009-0038-y.

Todt, E. and Torras, C. (2000). Detection of natural landmarks through multiscale opponent features, *15th International Conference on Pattern Recognition, ICPR-2000, Barcelona, Spain*, Vol. 3, pp. 976–979.

Wang, Q., Zhao, J., Huo, X., Wu, L., Yang, L. F.and Li, J. and Wang, J. (2018). Effects of a home care mobile app on the outcomes of discharged patients with a stoma: A randomised controlled trial, *Journal of Clinical Nursing* **27**(19–20): 3592–3602, DOI: 10.1111/jocn.14515.

Zawiślak, R., Mik, M., Mik-Wojtczak, J. and Granosik, G. (2020). Stoma-Alert—Development of IT system model supporting the process of diagnosing, treating and rehabilitating patients with stoma—Project no. POIR. 04.01.01-00-0066/18-01, *Technical report*, PHIN Consulting and Lodz University of Technology, Lodz.

Mariusz Szwoch received his MSc and PhD degrees in computer science from the Gdańsk University of Technology (GUT), Poland, in 1991 and 2002, respectively. He has been working at the GUT since 1991, and since 2019 he has been a university professor. His research interests include computer vision, image processing, artificial intelligence, 3D graphics, affective computing, and game development.

Rafał Zawiślak received his MSc and PhD degrees in automatic control and robotics from the Technical University of Lodz, Poland. For over 20 years he has been dealing with applications of computer science in medical systems. As part of his university work, he tackles control theory and numerical methods, but his real passion is designing database systems and friendly multimedia interfaces. He is a co-author of well-known and commonly used diagnostic systems in Poland.

Grzegorz Granosik was born in Poland in 1970. He received his MSc and PhD degrees in automatic control and robotics from the Technical University of Lodz, Poland. From 2002 to 2004, he worked as a postdoctoral researcher in the Mobile Robotics Laboratory at the University of Michigan, Ann Arbor. Since 2012 he has been a professor at the Lodz University of Technology, where he has founded the Human Centered Robotics Lab. His research interests include designing of novel robots and robot control systems.

Joanna Mik-Wojtczak has been involved in integrating business and science through innovative projects for more than 15 years. She is experienced in creating entrepreneurial environment for new business ideas that combine various fields of science. Since 2008 she has been actively developing PHIN Capital Group, an entity responsible for consulting activities in a wide spectrum of expertise. She received her MA in economics in 2003. Currently, she works on her doctoral dissertation on risk management in implementing R&D projects in medicine. The main areas of her interest are related to business and research: multidisciplinary R&D projects, risk management in R&D projects, development of innovation in enterprises, and project management.

Michał Mik is an assistant professor in the Department of General and Colorectal Surgery, Medical University of Lodz. He received his PhD and DSc degrees in 2008 and 2017, respectively. He is a specialist in general surgery and surgical oncology. For over 20 years he has been dealing with patients with gastrointestinal diseases including stoma patients.

Received: 15 April 2021
Revised: 30 July 2021
Re-revised: 7 September 2021
Accepted: 30 October 2021