Technique for automatic marker recognition augmented reality for images of architectural plans

N.A. Aksionova, A.V. Varuyeu, O.M. Demidenko, D.S. Sych
F. Skorina Gomel State University, Sovetskaya str, 104, Gomel, Belarus,
nataliaksen@gmail.com, varuyeu@gmail.com, demidenko@gsu.by, denis99sych@gmail.com

Abstract*. The article describes a technique for recognizing image markers of design and estimate documentation of the developer's architectural plans using computer vision technology. Existing off-the-shelf technologies and basic detection methods are being explored. The classification of methods according to the range of work is described. A recognition algorithm is proposed. The analysis of the Canny operator is carried out and its possible modification is considered to increase the efficiency of work when implemented in Python using OpenCV. Approbation of the developed software module is given.*

I. Introduction

The aim of the study is to find suitable technologies and algorithms for building a technique for recognizing augmented reality markers using computer vision methods. In the task set, square images of the developer's architectural plans are used as markers. When real-time image data is detected, a 3D model of the corresponding architectural structure is displayed on the screen of the mobile device. The main problem of marker recognition is to determine the marker in the video stream of the reader, set its location in the frame and project the virtual 3D model accordingly.

The authors of the article conducted a study of existing ready-made technologies and described the classification of methods according to the range of work. The article proposes an algorithm for recognizing markers and provides a possible modification of the existing Canny operator, where it is proposed to use the Rudin-Osher-Fatemi algorithm instead of Gaussian blur for image preprocessing. The proposed methodology will be used to implement applications using augmented reality technology.

The main class of label recognition problems using computer vision elements is associated with the detection of previously known markers. There are ready-made solutions for developing projects using augmented reality technology, such as the AR SDK. The most popular are Vuforia and EasyAR.

The recognition process depends not only on the quality of the marker and its location in space, but also on the distance of the reader's camera. In this regard, there is a classification of methods for detecting markers according to the operating range: near zone, middle zone, far zone.

The high resolution of the images used makes it possible not only to detect a label, but also to detect a sufficiently large number of input data, which ensures automatic reading of a significant amount of information. A significant disadvantage of the near-field methods is the use of large-sized images, which complicates the storage process and the duration of data processing. The study of the main methods for detecting augmented reality markers showed the relevance of developing our own SDK for the implementation of augmented reality applications.

II. Marker recognition algorithm

The marker recognition algorithm includes five main stages:

Step 1. Converting the input image to grayscale.

Step 2. Image binarization by adaptive thresholding.

Step 3. Determination of closed areas and selection of contours by the Canny method.

Step 4. Selecting the corners of the marker.

Step 5. Transformation of coordinates.

The process of determining the boundaries is based on determining the points of the digital image, where the brightness changes sharply. Canny's algorithm is based on calculating the image gradient, which shows the direction of the fastest increase of some value. The use of an edge detection algorithm simplifies image analysis while reducing the amount of data to be processed.

To implement the Canny algorithm, not only the search for gradient parameters is carried out, but also the primary image processing is performed: a Gaussian filter is applied to suppress noise in the image. Also, this algorithm uses the filtering of the obtained contours, which allows you to remove the so-called "extra" borders.

The main steps of the Canny algorithm are:

1. Converting a color image to grayscale.
2. Image smoothing by applying a Gaussian filter.
3. Search for gradients.
4. Suppression of "non-maxima".
5. Double threshold filtering.
6. Tracing the region of ambiguity.

The suppression of "non-maxima" is the process of translating the "smoothed" edges of the image of the gradient values into "sharp" edges. To do this, you need to compare the edge strength of the current pixel with the edge strength of the pixels in the positive and negative direction of the gradient. If the edge strength of the current pixel is greater than the edge strength of neighboring pixels, then the value remains, otherwise the value is reset.

Potential boundaries are determined by thresholds, among which the values of high and low thresholds are distinguished. After the stage of “non-maximum” suppression, the boundary pixels and the strengths of their faces will be identified by certain values, some of which will belong to the real edges of the image, and some of them will not determine the edges. To determine this circumstance, a threshold is determined and a double threshold filtering is performed. Border pixels with a strength value greater than the high threshold value are defined as "strong". Border pixels with a value below the low threshold value are removed, and those between the two thresholds are marked as “weak”.

III. Modification of the Canny algorithm

The use of Gaussian blur to suppress image noise in the Canny algorithm increases the stability of the results obtained after image preprocessing. However, this contributes to an increase in computational costs and even leads to loss of detail of borders and contours and distortion of the image itself. Gaussian blur can round the corners of objects and break edges at connection points.

For image preprocessing, the author of this article proposed the use of the Rudin-Osher-Fatemi (ROF) algorithm. This algorithm is used to clean the image from noise, while it smoothes the image, while maintaining edges and structures. This property is a weighty argument for using the ROF model, taking into account the specifics of the input images of design estimates, which are based on the contours of architectural plans. This article describes the mathematical model of ROF based on the Chambol algorithm.

The total variation of the halftone image *I* is the sum of the norms of the gradients, which in the continuous representation has the form:

$$J(I) = \int |\nabla I| dx. \tag{1}$$

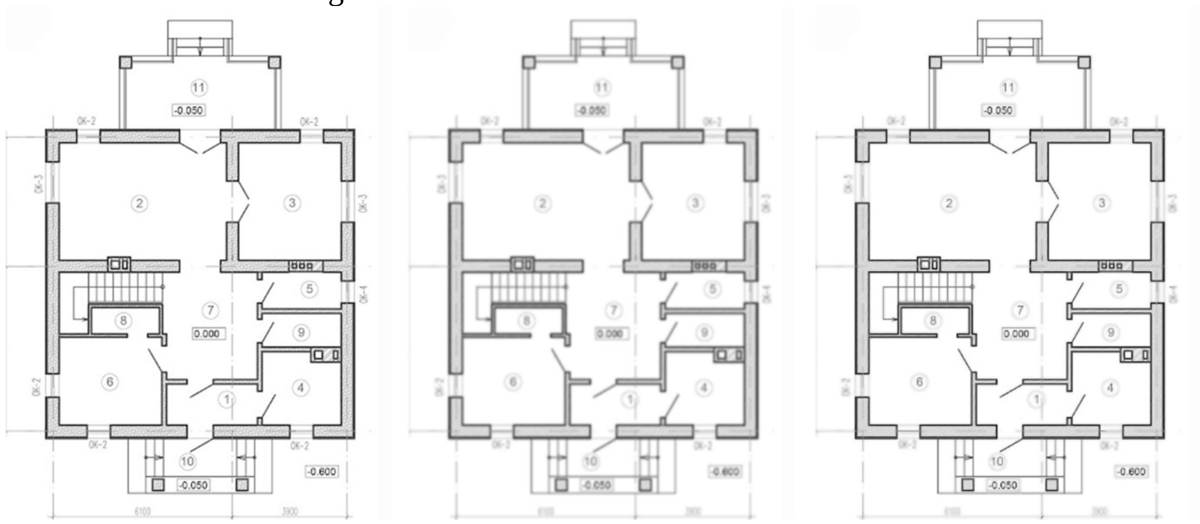
In discrete form, the formula takes the form:

$$J(I) = \sum_X |\nabla I|, \tag{2}$$

where summation is performed over all points of the image $X = [x, y]$. Chambol's variant of the ROF model finds a denoised image U for which the minimum of the function

$$\min_U \|I - U\|^2 + 2\lambda J(U), \tag{3}$$

where the norm $\|I - U\|$ measures the difference between U and the original image *I*. This means that this model is looking for an image that is "flat" but does not allow "jumps" at the boundaries between areas. Figure 1 shows the original image and the results of its processing using a Gaussian blur and a ROF model based on the Chambol algorithm.



a) b) c)

Fig. 1. Video stream frame processing algorithm
a) An example of denoising a grayscale image
b) Image after gaussian blur (σ=1.5)
c) Image after ROF model based on the Chambol algorithm

IV. Conclusion

As a result of the research, a method for determining markers for images of architectural plans was proposed, and a recognition algorithm was described. Also, the authors proposed a modification of the Canny operator, where the Rudin-Osher-Fatemi algorithm is used for image preprocessing instead of Gaussian blur. Approbation of the modified Canny algorithm showed high efficiency in extracting contours, while maintaining boundaries and structures. The software add-on for defining markers was implemented in Python using the OpenCV library.

References

[1] *N. Aksionova.* Development of SDK for a mobile application using augmented reality technology // N.A.Aksionova, A.I.Kucharau / Bulletin of F. Skorina Gomel State University, No. 3 (126) pp 81-84 (In Russian)

[2] *N. Aksionova*. Presentation and implementation of educational material using augmented reality technology // N.A.Aksionova, A.I.Kucharau / Trends and prospects for the development of science and education in the context of globalization (International scientific conference. Electronic materials) No. 71 pp 235-239

[3] *Demidenko OM, Aksionova NA.* Development of a Machine Vision System for Image Recognition of Design Estimates. Nonlinear Phenomena in Complex Systems. 2022; 25(2):159 – 167. DOI: 10.33581/1561-4085-2022-25-2-159-167.

[4] *Demidenko OM, Aksionova NA, Varuyeu AV and Kucharav AI*. 3D-modeling of Augmented Reality objects using Shi-Tomasi corner detection algorithms. J. Phys. 2021;2091 (012058). DOI: 10.1088/1742-6596/2091/1/012058