

Automated Computer Vision-Based Reading of Residential Meters

Karlo Koščević

Faculty of Electrical Engineering
and Computing
University of Zagreb
Email: karlo.koscevic@fer.hr

Marko Subašić

Faculty of Electrical Engineering
and Computing
University of Zagreb
Email: marko.subasic@fer.hr

Abstract—We present a solution for automated reading of various residential meters from photographs, using computer vision algorithm. The solution has several modules that are executed sequentially: meter type recognition, geometric transform of images, ROI extraction and OCR. Algorithm uses SURF and HOG features to detect and describe feature points used for device recognition and OCR. The final goal is to provide complete counter values and complete serial number of the meter. The solution allows for a limited amounts of poor imaging conditions, and usually fails in poor image conditions when even human observers would have difficulties reading meters. The solution has been implemented in MATLAB environment and its computer vision library. An initial image database has been collected for testing purposes. Test results are reported in the paper.

I. INTRODUCTION

The goal of presented research is to enable automated reading of relevant data from images of residential meters, such as one shown in Figure 1. The solution is intended and applied to meters with mechanical counters, but the proposed method could easily be adapted to meters with electronic displays. The proposed solution should effectively turns ordinary residential meters into automated smart meters at no extra cost. The images of meters could be obtained using a smartphone, which could also be use for processing of the image. Ideally, a meter should be photographed at a right angle, with sufficient light, good contrast and no reflections, but the proposed method is capable of compensating for a limited amount of imperfections in images. However, certain light constraints regarding image acquisition must be obeyed.

The problem of automated extraction of visual information may seem trivial to an average person, because humans can easily detect all relevant areas of a meter, and also read out relevant information. A computer vision algorithm that does the same, must have several distinct steps. The steps are: detection and recognition of the specific meter in the image, geometric correction of the image, extraction of relevant regions of interest from the image, and actual reading of numerical data. The algorithm extracts two types of regions of interest, one or more counters and also an ID number, usually a serial number of the meter. Once the numerical data has been successfully read, it can easily be forwarded to the corresponding utility company, and used for personal record of the user.



Fig. 1: Example of photographed device

Device detection, recognition and reading of numerical characters is performed by template matching, so appropriate template database had to be prepared for the set of residential meters used in this research. If reading of an additional type of meter is required, new templates should be added to the database.

II. METHODS USED

Several image processing and analysis tools have been used in the proposed solution so each is briefly described in this section.

A. SURF

Speed Up Robust Features or shortly SURF [1] is both detector and descriptor of image features. It is partly based on SIFT descriptor [2], but its authors claim that SURF is much faster and robust. It can be used for tasks such as image registration, object recognition, classification or 3D reconstruction. For detection of significant feature points, SURF uses a blob detector based on Hessian matrix. Hessian matrix is calculated at each image pixel I at coordinates p q as shown in Eq. (1). Also, in improve performances, integral images are used. The determinant of Hessian matrix represents a measure of local change around points and points where this determinant form local maximum are chosen. Feature points are looked for at different scales with initial box filter size of 9×9 (corresponding to Gaussian derivatives with $\sigma=1.2$).

Invariance to rotation and scaling are also useful qualities of SURF.

As for the local neighborhood descriptor, it consists of fixing a reproducible orientation based around the interest point and extraction of SURF descriptor from a square region aligned to the assigned orientation. The orientation is calculated in order to archive rotational invariance of feature vector. The dominant orientation is estimated by calculating the sum of all Haar wavelet responses weighted by a Gaussian function within the sliding window of size $\pi/3$. Region around the point is described by extraction of square region centered on the point and oriented along the selected orientation. For each sub-region of extracted square (that is split into 4x4 square sub-regions) Haar wavelet response is extracted and weighted with Gaussian. Gaussian filter affects the features in the manner that makes them more robust to the noise, translation and deformations.

$$H(p, q) = \begin{bmatrix} I_{xx}(p, q) & I_{xy}(p, q) \\ I_{xy}(p, q) & I_{yy}(p, q) \end{bmatrix} \quad (1)$$

B. HOG

The Histogram of Oriented Gradients or shortly HOG is a image feature descriptor [3]. It is widely used for purposes of object detection. As its name points, HOG is based on computation of gradient orientations. Image is divided in smaller regions called blocks and each block is also divided in smaller regions called cells. Then in each cell, gradient amplitude and orientation are calculated, and cell is assigned to appropriate orientation bin. Once that is done for all cells in one block, all this data from cells is normalized and concatenated into one vector that represents a descriptor of single block. Afterwards, block descriptors are concatenated and they form a single vector that is a descriptor for the whole image. Because of this principle, HOG is also called a global descriptor and the size of its vector depends on number of blocks and cells, as well as number of orientation bins. It is important note that HOG is not invariant to image rotation.

C. Perspective projection and Homography

Perspective projection is a geometric transform of points form 3D scene to the 2D plane. This happens regularly during 2D image acquisition when projection is effectively done at an image sensor plane. Basic concept of this projection is presented in Figure 2. In our case, we will be comparing two projections (photographs) of the same object projected at different planes. A perspective transform or homography [4] will be required to transform input image to the projection plane of the template image. This will effectively align meter screens of input images to meter screens of template images. Here we presume that meter screens are entirely flat, which is not entirely true. To estimate homography parameters, at least four corresponding point pairs are required, but more can be used to obtain a more reliable estimate.

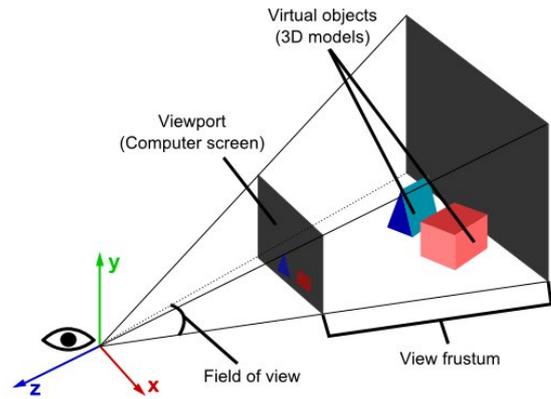


Fig. 2: Example of perspective projection



Fig. 3: Samples of residential meters

D. K-d tree

K-d tree is a special case of binary space partitioning trees [5]. K-d trees are arranging points in k-dimensional space and are useful data structure for multidimensional search such as range searches and nearest neighbor searches. The algorithm works by recursively partitioning the set of training instances based on a median value of a chosen attribute. Attributes are alternately selected coordinate axes of training instances. When we get a new data instance, we find the matching leaf of the K-d tree (alternately comparing its data with nodes' data), and compare the instance to all the training point in that leaf. We use k-d tree to increase speed of feature point matching process (searching of two nearest feature points to the one that is observed).

III. INPUT DATA AND TEMPLATE DATABASE

We have obtained 79 images of residential meters for electricity and gas (several samples showed in Figure 3). Though we had just two meter categories, differences in positions of counters and serial numbers required separate templates with different ROI positions for the same meter type. Differences in screen color, digit color and some larger details, also required separate templates, as otherwise, chances of false



(a) Template image before homography. Green circle - initial corners; white dashed line - initial x and y coordinates of initial corners; solid red line - new x and y coordinates of corners. Each corner is translated to its closest intersection of two red lines.



(b) Template image after homography

Fig. 4: Illustration of homography of meter template

meter recognition are quite high. Hence, effectively we had 9 meter classes, and we had prepared 9 templates.

A template consists of cropped image of the meter screen, with screen rectified (see example in Figure 4). In order to perform cropping and rectification a homography has to be performed. Here we presume that meter screen is completely flat and all counter digits lie in the same plane. This is not entirely true for common residential meters, but distortions induced by this step are negligible in our expected usage scenario, where template is produced using the best obtained image for given meter type (least amount of blurring, noise, no saturation, camera positioned at the right angle). To estimate appropriate homography parameters, four corners of template screens have been manually annotated. Together with desired corner positions of templates, they give all homography parameters. After applying homography, all meter screen edges should be at the right angle as illustrated in Figure 4.

For the proposed method, regions of interest (ROI) information has to be stored with each template. Coordinates of two opposite corners (of rectangular ROIs) in template coordinate space are stored for each counter digit and for serial number. The coordinates are annotated manually in the rectified template image.

SURF feature points are calculated for each template image. Obtained feature points are filtered to remove feature points outside borders of the meter screen. Feature points are also removed if they are positioned in or very near to the variable

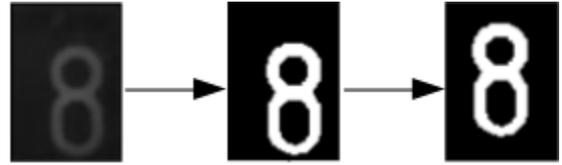


Fig. 5: Preparation of sample image of digit

regions of meter screens. Such variable regions are counters and serial numbers and maybe some minor variable details in meter screens. The idea is to obtain significant feature points that are highly likely to appear with all meters of the given type.

Also, a single binary template image has been prepared for each digit. We haven't noticed significant differences in number fonts of counters and serial numbers between different meter classes. Hence, just ten number templates (one for each numerical character) has been used in this research. Digit templates have also been chosen with the image quality in mind, and each digit template is of the same size 60x84 pixels. Binarization is performed by thresholding, using average gray value of the template as a threshold. Binary representation of the digit is placed centrally within the template image (see example in Figure 5). Binary templates will be compared with binarized digits from the input image.

Our testing database consisted of 3 to 35 images per each meter class. Images were obtained by smartphone cameras in various conditions, and image dimensions were 720 to 3648 x 776 to 3648 pixels. Corner points of all meter screens have been manually annotated for validation purposes.

IV. OVERVIEW OF THE PRESENTED METHOD

The presented method starts with recognition of the residential meter in an image. SURF is used for that purpose which makes recognition process insensitive to translation, scaling, image rotation, and a limited amount of in-plane rotation. This enables certain degree of flexibility regarding the position of the camera relative to the screen of the meter. Exaggerations in geometric transformations usually makes meters unreadable even for human readers, so our solution focuses only on images with moderate amounts of geometric transforms. We also assume that whole screen of a meter is visible in an image, and that the screen spans most of at least one image dimension. Such requirements should easily be satisfied if a person is willingly photographing a residential meter. Certain cases when these conditions are not satisfied, could even be detected automatically in real-time to warn the user. Expected image conditions are also reducing the problem of meter detection, as we expect that just one meter screen is present in the image.

The recognition is performed by matching of the SURF points detected in an image with SURF points of all templates. The right meter template should have the largest number of matched points. After successful recognition, the pairs of matched SURF points are used to estimate geometric

transform needed to align the meter in the image with the meter in the template. Once the image is aligned with the template, ROI positions of the template can be used to extract ROIs from the image.

The final step is reading the numerical data for extracted ROIs containing counters and serial number. The appearance of these two numbers usually differ, so slightly different procedures have to be applied. Each counter digit is usually clearly physically separated from others, and all digits can use same size bounding box. This makes counter digit extraction an easy task. Serial numbers are printed, and usually characters don't have fixed widths. Hence, prior to character recognition, serial number has to be segmented. Numerical characters are finally recognized by comparing them to character templates using HOG features.

Significant error in any of the above steps, increases the chance of wrong reading results or inability to properly read data from the meter. Failure to detect an expected character in the ROI can be detected, and the user warned.

V. RECOGNITION OF RESIDENTIAL METERS

First step of meter recognition is SURF feature point matching. Each significant SURF feature point in an input image is matched against all SURF feature points of all templates. Matching is done by searching for the nearest neighbor in k-d tree which is made of feature points extracted from all templates. For each feature point of an input image, two nearest neighbors among template feature points in the feature space are obtained. The Euclidean distances of two nearest neighbor feature points to the input feature point have to be sufficiently different to avoid confusion and to guaranty successful match. If the difference is sufficient, the input feature point can be matched to its closest neighbor. If the difference is below a certain threshold, the input feature point is not matched. Finally, the match is accepted if the distance is below a threshold, otherwise the match is discarded. Once all feature points from an input image have been processed, each will have one or zero matches. An example of matched points for one template can be seen in Figure 6.

We expect that a meter screen in an input image can be adequately aligned with the screen in an template image by homography. All coordinate differences of matched feature points between input image and an template, must correspond to the same transform parameters. Hence, the matched pairs are used to estimate parameters of the transform. If coordinate difference of any matched pair does not conform to the estimated geometric transform, the match is discarded. Since each matched pair presents a vote for a corresponding template, this step reduces the probability of that template match.

An input image is matched with a template that has the largest set of remaining matching feature points.

VI. ROI EXTRACTION

Once the meter has been successfully recognized, rectangular ROIs containing relevant numerical data have to be extracted. Each template has a set of ROI coordinates in its

coordinate system, so meter screen in the input image has to be aligned with meter screen in the template image. We presume that a meter screen together with its whole content is flat so a homography transform has to be performed to align two meter screens. An input image will not violate the assumption if it has been recorded according to the acquisition requirements of the proposed method explained in section III.

Parameters of the geometric transform have to be estimated using a set of matching points in two images. However, matching pairs of significant points have already been obtained in the previous step. Furthermore, parameters of the geometric transform have also been obtained for the matching refinement. Hence all that is required in this step is to apply the geometric transform to the input image. One requirement is that there is a sufficient number of matched pairs. Four matched pairs is the minimum, but more pairs is preferred. Low number of matched pairs is usually indicator of false match or poor imaging conditions. An example of the geometric transform is shown in Figure 7. One can observe that in the transformed image, meter screen is in the upright position with all four edges at the right angle.

To evaluate the quality of geometric transform, corners of each meter screen have been manually annotated. After the input image transform, new coordinates of screen corners can be compared to the corner positions of the corresponding template. Results section contains evaluation results for geometric transformations.

Once meter screens are aligned, ROI coordinates of the template are used to extract each counter digit and a serial number.

VII. OCR

Digits are recognized using HOG descriptors. We have chosen HOG for this purpose for the reason that the edges of objects that are recognized can be clearly separated and the assumption is that objects can differ based on finite number of cells and orientation bins. Rectangular ROI containing a digit is extracted in the previous step, so the recognition is performed in the ROI only. Each digit is preprocessed in the same way as digit templates to obtain a centered binary representation of the digit. The binary image of the digit must be scaled to match the size of digit templates. HOG vector is calculated for the scaled digit image, and the scaling ensures that HOG vectors of input image and templates have the same size. Images are divided into cells so that each cell has both width and height equal to a quarter of the image width and height. The cell histogram is 1-by-*NumBins* vector where *NumBins* is set to a default size of 9 orientation histogram bins. Blocks consists of 4 cells with 1-by-1 cell overlap between blocks. Total number of blocks is 12 and total size of calculated HOG vector is 1-by-216. A digit is recognized as a number whose template has the nearest HOG vector. There are only ten template vectors to compare with moderate number of elements, so no optimized comparison algorithms are required (like k-d tree in section V).

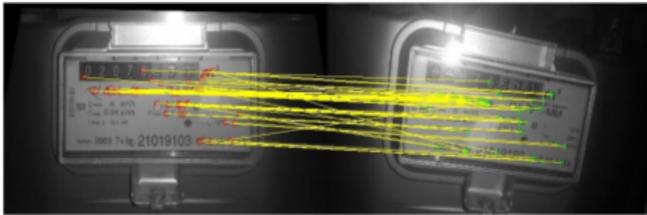


Fig. 6: Visualization of matched features. On the left side is the sample device to which observed one (one on the right) was matched.



(a) Meter before the projection



(b) Meter after the projection

Fig. 7: Visualization of the projected meter

This procedure can be applied directly to counter digits, where each digit has its own bounding box, and bounding boxes are clearly separated. Due to this a separate ROI for each digit is stored with each template. Digits in serial numbers have varying widths (e.g. digit 1 is narrower than other digits) and are not equally spaced. A segmentation step is required to separate digits after the entire serial number has been successfully extracted. The serial number ROI is binarized using its average gray value as a threshold, and sufficiently large connected components are extracted as separate digits. Each digit is then represented as one image and dimensions of the image are equal to width and height of connected component which is assumed to be a digit. Images of digits are once more binarized and HOG vector is calculated for each digit and compared to HOG vectors of digit templates.

TABLE I: Test results of reading residential meters

Device type	No. of images	Correct recognition	Correct geom. transf.	Correct reading of counter (%)	Correct reading of serial No. (%)
1	7	7	7	89.8	94.64
2	35	35	33	96.33	85.71
3	3	3	3	81.95	75
4	5	3	1	100	-
5	20	19	16	86.785	90.15
6	1	1	1	100	88
7	3	3	3	86	83.33
8	5	4	2	100	87.5
TOTAL	79	75	66	92.61	86.19



(a) Image with high level of noise



(b) Device captured from bad angle

Fig. 8: Example of bad images for process of reading residential meters

VIII. RESULTS

As it is explained in section V, to match two features, two thresholds have to be applied. The Euclidean distances of the first nearest neighbor feature point to the input feature point has to be at least 22.5% smaller than the distance of the second nearest neighbor feature point. If the input feature has been matched to its closest neighbor, the match is accepted if their Euclidean distance is below value of 0.25, otherwise the match is discarded.

Results for each algorithm step and each device class are shown in Table 1. Images which were producing an error in one of the steps of algorithm were discarded after that step. Columns annotated as *Correct reading of counter* and *Correct reading of serial number* are representing percentage of meters on which each digit of counter or serial number was correctly read.

Geometric transform of an image was declared unsuccessful when number of matching pairs between observed image and a template image was too small (5 or less matches) or estimated transform was very unlikely. If the number of matching pairs was sufficient, the recognition was always successful. As for reading the serial number, wrong recognitions were mostly due to errors in binarization and the detection of connected components. Usually, the threshold value calculated as the average intensity of the ROI would produce inadequate connected components. This would usually be caused by poor imaging conditions. Examples of bad images of device are shown in Figure 8. Subfigure 8a has excessive amount of shadows and background reflection, as well as the reflection of light which is the major problem in most cases. Although, SURF is invariant to rotation, angle of device in subfigure 8b is too large and after homography, some parts of digits can not be correctly determined. Clearly, our approximate assumption that meter screens are flat, does not hold in this case.

IX. CONCLUSION

SURF and HOG have proved to be appropriate tools for recognition of residential meters and reading of their numerical data. Although, the results for reading digits are slightly lower than ones for recognition of device type, complete algorithm has overall very good results, and it can be applied on many different types of devices displaying numerical data. All that needs to be done to read devices of a new type is to prepare one or several templates that represent that device type. Algorithm certainly has room for future improvements, especially part of the algorithm that segments the serial number. Most errors occurred in this procedure, while the other parts had success rate around 90% or greater (recognition of device type had a success rate of 95%). After examining the results it can be seen that the most common errors are caused due to the poor quality images, e.g. images with high level of reflected background light, images with shadows or blurred images. Therefore, images should be obtained carefully to assure sufficient image quality. For this reason, our future plans include automatic estimation of image quality.

REFERENCES

- [1] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. *SURF: Speeded Up Robust Features*, pages 404–417. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.
- [2] D. G. Lowe. Object recognition from local scale-invariant features. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1150–1157 vol.2, 1999.
- [3] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01*, CVPR '05, pages 886–893, Washington, DC, USA, 2005. IEEE Computer Society.
- [4] R. Baer. *Linear Algebra and Projective Geometry*. Dover Books on Mathematics. Dover Publications, 2005.
- [5] Jon Louis Bentley. Multidimensional binary search trees used for associative searching. *Commun. ACM*, 18(9):509–517, September 1975.