A SET OF DEFINITIONS FOR WORKING WITH SPATIAL FILTERS

Luděk Bryan
Doctoral Degree Programme (5), FIT BUT
E-mail: crha@fit.vutbr.cz

Supervised by: Vladimír Drábek
E-mail: drabek@fit.vutbr.cz

ABSTRACT

This article sugguest a new set of mathematical definitions suitable for description of spatial filter operations. First, there are definitions of image, pixel, structuring element and neighborhood. Based on those, definition of filter and filtered image follows. As an example of usage of those definitions, preprocessing and feature extraction filter of the method proposed in my PhD thesis is suggested.

1 INTRODUCTION

My PhD thesis deals with a new object detection method. This method is based mainly on spatial filters. However, while studying relevant literature sources, I didn't find any suitable set of mathematical definitions for working with spatial filters, or generally with other computer vision techniques. For example, [1] uses algorithmic style of description which already implies implementation style. [2] uses mathematical style, but many things are left undefined thus definitions are rather informative than comprehensive. Similar style to that one in [2] is also used in many articles that are dealing with computer vision.

This article suggests a new set of mathematical definitions suitable for description of spatial filter opeprations. As an example, the method for object detection suggested in my PhD thesis is described using those definitions.

2 THE DEFINITION SET

2.1 IMAGE

Image (Definition 2.1) is a matrix of size $w \times h$ consisting of *pixels* (picture elements). Each pixel is defined by its color. For our purposes, we will consider only gray-scale images, where the only used color is gray ranging between black and white. Thus, pixel can be described by one integer value between 0 (black) and $d-1$ (white), where d is the number of gray levels, called *gray depth*.

Definition 2.1 *Image G of size $w \times h$ and gray depth d is a matrix*

$$G = \begin{pmatrix} g_{00} & g_{10} & \cdots & g_{w0} \\ g_{01} & g_{11} & \cdots & g_{w1} \\ \vdots & \vdots & \ddots & \vdots \\ g_{0h} & g_{1h} & \cdots & g_{wh} \end{pmatrix}$$

where $g_{ij} \in \{0,1,\ldots,d-1\} \forall i \in \{0,\ldots,w-1\}, j \in \{0,\ldots,h-1\}$

Value g_{ij} (pixel)of image G can be also written as $G_{i,j}$.

For easier notation, we also define the set of all images with certain width, height and gray depth $\mathcal{G}^d_{w,h}$, images with certain width and height $\mathcal{G}_{w,h}$ and set of all images $\mathcal{G}$ in definition 2.2.

Definition 2.2 *Let us denote the set of all images of size $w \times h$ and gray depth d*

$$\mathcal{G}^d_{w,h}$$

The set of all images of size $w \times h$ is

$$\mathcal{G}_{w,h} = \bigcup_{d \in \mathbb{N}} \mathcal{G}^d_{w,h}$$

and the set of all images is

$$\mathcal{G} = \bigcup_{\substack{w \in \mathbb{N} \\ h \in \mathbb{N}}} \mathcal{G}_{w,h}$$

Sometimes, we need to refer to a certain pixel. This is done through *coordinates* (x,y) (definition 2.3), that are pointing to a pixel $G_{x,y}$.

Definition 2.3 *The coordinates of a pixel of an image $G \in \mathcal{G}_{w,h}$ is an ordered pair (x,y) s.t. $x \in \{0..w-1\}$ $y \in \{0..h-1\}$*

2.2 STRUCTURING ELEMENT

As we already said, filter is applied to the neighborhood of a pixel. For each pixel that we apply filter to, neighborhood has the same shape. Thus, we need to define the shape of the neighborhood we are applying filter to. This shape will be called *structuring element*. We will define structuring element as a set of ordered pairs of integer numbers (definition 2.4). These ordered pairs will then be used as relative coordinates to the filtered pixel.

Definition 2.4 *Structuring element S is a set of ordered pairs of integer numbers, i.e.*

$$S \subseteq \mathbb{Z} \times \mathbb{Z}$$

Again, we will also define the set of all structuring elements (definition 2.5).

Definition 2.5 *Let $\mathcal{S}$ be the set of all structuring elements.*

2.3 PIXEL NEIGHBORHOOD

Now, we can define actual *neighorhood* of a pixel covered by a structuring element (definition 2.6). Neighborhood contains actual pixels that are used as source values for a filter function.

Definition 2.6 *Neighborhood N of a pixel at coordinates (x,y) in image $G \in \mathcal{G}_{w,h}$ with structuring element $S \in \mathcal{S}$ is a matrix*

$$N = \begin{pmatrix} n_{00} & n_{10} & \cdots & n_{w0} \\ n_{01} & n_{11} & \cdots & n_{w1} \\ \vdots & \vdots & \ddots & \vdots \\ n_{0h} & n_{1h} & \cdots & n_{wh} \end{pmatrix}$$

$$\begin{aligned} \textit{where } n_{x'y'} &= G_{x',y'} \forall (x',y') \ s.t. \ \exists (i,j) \in S \ \ x' = i + x, \\ &\qquad\qquad\qquad\qquad\qquad\quad y' = j + y, \\ n_{x'y'} &= \perp \ \ \textit{otherwise} \end{aligned}$$

Value n_{ij} of neighborhood N can be also written as $N_{i,j}$.

Definition 2.7 *The set of all neighborhoods on image G at coordinates (x,y) with structuring element S is denoted*

$$\mathcal{N}^{G,S}_{x,y}$$

The set of all neighborhoods with structuring element S is

$$\mathcal{N}^{S} = \bigcup_{\substack{G \in \mathcal{G} \\ x \in \{0..w-1\} \\ y \in \{0..h-1\}}} \mathcal{N}^{G,S}_{x,y}$$

The set of all neighborhoods is

$$\mathcal{N} = \bigcup_{S \in \mathcal{S}} \mathcal{N}^{S}$$

where w, h are width and height of an image G.

2.4 FILTER

Finally, we can define a filter in definition 2.8. Filter transforms neighborhood at any image at any coordinates, but with defined shape (structuring element) to a real number. Thus, while defining a filter function, structuring element is known, but the filter can be applied to any pixel at any image. Actual filter function F is not specified here, its variants are described more in detail in the following subsections.

Definition 2.8 *Filter of structuring element S is a function $F : \mathcal{N}^{S} \rightarrow \mathbb{R}$*

Definition 2.9 *The set of all filters of structuring element S is denoted*

$$\mathcal{F}_S$$

The filter is supposed to be applied to all pixels in an image. This is shown in definition 2.10, where G and G' are input image and filtered image, respectively, and S is a structuring element of a filter.

Definition 2.10 *Let $G \in \mathcal{G}_{w,h}$ be an image, $F \in \mathcal{F}_S$ a filter. Filtered image $G' \in \mathcal{G}_{w,h}$ filtered by F of source image G is*

$$G'_{x,y} = F(N)\ s.t.\ N \in \mathcal{N}_{x,y}^{G,S}\ \forall x \in \{0..w-1\}, \forall y \in \{0..h-1\}$$

3 OBJECT DETECTION METHOD

The basic idea of the proposed method is to implement a large number of search units, each searching for one fragment of the object. During implementation, the standard sequence of steps for pattern recognition should be used - preprocessing, feature extraction, and classification. Example of real-life application, detection of a license plate, is shown in Figure 1. Mathematical definitions of preprocessing and feature extraction are following.

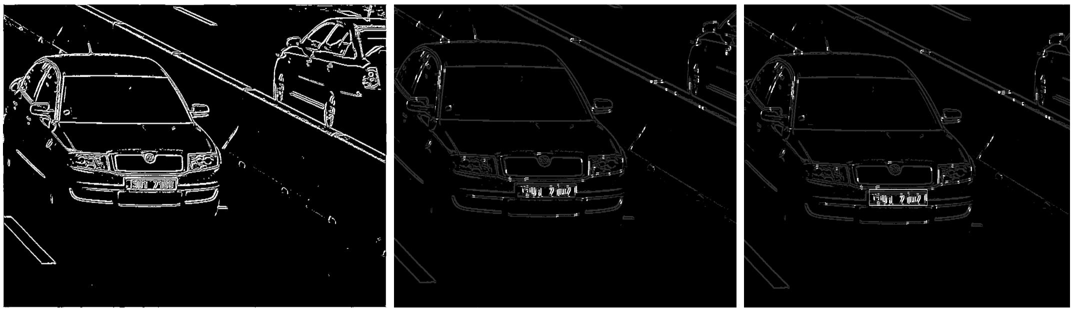

Figure 1: Example of preprocessing and feature extraction and classification

3.1 PREPROCESSING

The preprocessing unit uses a local thresholding technique. The whole operation is implemented as a filter. First, a threshold value called the *divider* is computed for each pixel from its neighborhood. It is a function of the neighborhood pixel values like median, average, etc. The filter then compares this value to the active pixel, and outputs either 0 for dark or 1 for bright pixels (Definition 3.1). The input to the preprocessing unit is an image $G \in \mathcal{G}_{w,h}^d$ and the output is an image of the same size $G^p \in \mathcal{G}_{w,h}^1$, but with gray depth of 1.

Definition 3.1 *Let $F \in \mathcal{F}_S$ be a divider filter. Preprocessing filter $P \in \mathcal{F}_S$ of filter F is defined as follows.*
Let $N \in \mathcal{N}_{x,y}^{G,S}\ \forall G \in \mathcal{G}_{w,h}\ \forall w,h \in \mathbb{N}\ \forall x \in \{0..w-1\} \forall y \in \{0..h-1\}$. Then

$$P(N) = \begin{cases} 1\ F(N) < G_{x,y} \\ 0\ otherwise \end{cases}$$

3.2 FEATURE EXTRACTION FILTER

The feature extraction filter is the core of my method. It is designed to find certain components in an image, i.e. basic shapes, letters or parts of letters. The input to the feature extraction unit is the output image from the preprocessing unit G^p. The output is an image of the same size containing information about matching templates $G^f \in \mathcal{G}_{w,h}^{d_f}$.
To implement the feature extraction unit, first we need a data structure to store information about the components we are searching for. We will store them as small binary images called *templates*. The set of templates belonging to one object will be called the *template bank*.
Then we need to create a filter that can detect a template. Basically, we need to *compare* similarities between two images (a template and a slice of an image). This leads us towards using a comparison operation in the filter. According to experiments, the most suitable operation for hardware implementation seems to be a *hit-or-miss* filter, that outputs 1 whenever an image slice and a template match. The feature extraction filter will then be a set of filters, each searching for one template.

Definition 3.2 *Template for structuring element S is function* $T : S \rightarrow \{0, 1\}$

Definition 3.3 *Set of all templates for structuring elements S is denoted* $\mathcal{T}_S$

Definition 3.4 *Template bank B for structuring element S is function*

$$B : \mathcal{T}_S \rightarrow R_\perp$$

Definition 3.5 *Set of all template banks for structuring elements S is denoted* $\mathcal{B}_S$

Definition 3.6 *Let* $N \in \mathcal{N}_{x,y}^{G,S}$ *be a neighborhood for some* $G \in \mathcal{G}_{w,h}^2$ $w, h \in \mathbb{N}$ $x \in \{0..w-1\} y \in \{0..h-1\}$ *and* $T \in \mathcal{T}_S$ *a template. Operator* $\approx$ *(match) is defined as follows:*

$$N \approx T \Leftrightarrow \forall i, j \in S : N_{x+i,y+j} = T(i, j)$$

Feature extraction filter $F \in \mathcal{F}_S$ *for template bank* $B \in \mathcal{B}_S$ *is every filter that satisfies the following condition:*

$$F(N) \neq \perp \Rightarrow \exists T \in \mathcal{T}_S \ s.t. \ B(T) = F(N) \land T \approx N$$

4 CONCLUSIONS

My PhD thesis deals with a new object detection method. This method is based mainly on spatial filters. In this article, I suggested the new set of definitions for working with spatial filters. As an example, preprocessing and feature extraction filters, the core of my PhD thesis, have been described using the suggested definitions.

REFERENCES

[1] Davies, E.R.: Machine Vision: Theory, Algorithms, Practicalities, 3rd edition, San Francisco, CA, USA, Elsevier inc., 2005, ISBN 0-12-206093-8

[2] Jahne B., Hausecker H., Geisler P.: Handbook of Computer Vision and Applications, Academic Press, San Diego 1999, ISBN 0-12-379770-5