# Layered RC Circuit Model for Background Subtraction

Karel Mozdřeň, Eduard Sojka, Radovan Fusek, and Milan Šurkala

Technical University of Ostrava, FEECS, Department of Computer Science,
17. listopadu 15, 708 33 Ostrava-Poruba, Czech Republic
{karel.mozdren, eduard.sojka, radovan.fusek, milan.surkala}@vsb.cz

**Abstract.** The background subtraction is a technique widely used for video analysis, mainly moving object detection for surveillance systems. Such algorithms must be robust, fast and it has to be able to deal with dynamic backgrounds like water surface or moving tree branches. Also, they should be able to deal with illumination changes and objects casted shadows. Generally, in computer vision the algorithms with a physical background have the best performance. We propose an algorithm for background subtraction based on a model of layered RC circuits. We tested our method on video sequences acquired from level crossing and on commonly used datasets. Finally, we have compared the proposed method with other frequently used methods.

## 1 Introduction

The background subtraction is a common technique for moving object segmentation from video sequences. It is an alternative to the object detectors based on the knowledge of appearance of the objects. There are two popular methods: Ada-boost proposed by Yoav Freund and Robert E. Schapire [1], and support vector machine by Constantine Papageorgiou and Tomaso Poggio [2]. Both are trainable object detectors and are used in variety of applications.

In many cases, we are not able to predict the size, shape, or color of the objects, tahtin we are trying to detect. In such cases, we are ought to use the background subtraction algorithms. These, instead of training the object detector, model the background and subtract from it the *"moving"* objects in the foreground. The result of background subtraction is a binary image, where the pixels indicating *"moving"* objects are marked with white color and background pixels with black color. Then the connected components algorithm is used to find the individual objects in the segmentation. An example application of such an algorithm is the project Pfinder by Christopher Richard Wren et al. [3]. Their system tracks people and interprets their behavior. The tracking itself is done using the background subtraction algorithm. Generally, PDE based algorithms have the best performance in computer vision applications. Our method is inspired by the model proposed by Pietro Perona and Jitendra Malik [4]. They presented a model, using a grid of resistors and condensers organized in a grid, functioning as an image filter. In this paper we present our modification to this model. We have modified the filter so it can be used for background modeling. The model is extended by excitation contacts used for connection of the input video sequence images with the model. Our paper has the following structure: in Section 2 we present works related to our matter, in Section 3 we describe our modified model, in Section 4 we conduct experiments and compare our method with other commonly used methods, and in Section 5 we conclude our work.

## 2 Related Works

The last two decades witnessed great improvement in background subtraction algorithms. Background subtraction techniques are based on learning of the background model from the video sequences. Each image in the video sequence is subtracted from the background model to find the differences and is also used for adaptation (update) of the background model. Big difference between the model and the actual image indicates *"moving"* objects in the foreground. One of the earliest method was proposed by Alan Lipton Hironobu et at. [5]. We refer to it as a temporal difference background subtraction (TDBS). This method uses one of the resent images as the background model and subtracts it with the most recent image. The difference is then thresholded and objects in the foreground are found. This method is not prone to fast illumination changes, but is prone to slow or temporally stationary objects. When the background model is too similar to the input image, there is too little difference and no foreground objects are indicated. This is exactly what happens when slow objects are moving in the images. Another method was proposed by Christopher Richard Wren et al. [3] and is known as temporal Gaussian background subtraction (TGBS). This method creates statistical model of the background. Each pixel of the background model is represented by one Gaussian defined by two values: $\mu$ for the mean value and $\sigma$ for the standard deviation, both computed from $N$ recent images. For better performance, those values are computed using the running Gaussian, which approximates these two values and is not that much demanding upon memory and computational time. This method deals with the problem of slow or temporally stationary objects detection, because it uses more than one image for background model construction. As a trade off, it is prone to the illumination changes. Sometimes, the $\mu$ value is represented by median (TMBS) [6], it makes the method more stable if the time between individual video sequence images vary, but the computational time and memory requirements rise accordingly. The problem with all these methods rises with dynamic backgrounds. Dynamic backgrounds are for example: moving tree branches, water surface waves, etc. . The method dealing with dynamic backgrounds is known as the mixture of Gaussians background subtraction (MoGBS). It was developed by C. Stauffer and W.E.L. Grimson [7]. This method models not only one, but $K$ Gaussians for each pixel, and each Gaussian adapts to one background. For example, in the case of moving branches, one background represents the branch and the other the sky behind it. This also applies for interior video sequences, where the light is switched on and off. In this case, one Gaussian adapts to dark background and the other one to the illuminated background.

We describe our method in the next Section. First, we show our modification to the model presented by Pietro Perona and Jitendra Malik, then we propose a layered version of the model for dynamic backgrounds, and finally we provide the reader with information about additional improvements to our method.

## 3 Proposed Method

The inspiration to use a layered RC circuit model for background subtraction came from paper by Pietro Perona and Jitendra Malik [4]. In their work, they presented a method for scale-space image filtering based on heterogeneous isotropic diffusion. They also presented a model of diffusion process using electrical components, namely resistors and condensers (RC circuit). The scheme for filter they proposed can be seen in Figure 1. As we can see, the circuit is a grid of

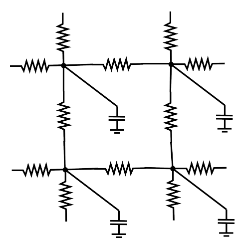


**Fig. 1.** The image filter model proposed by Pietro Perona and Jitendra Malik

condensers connected to neighboring condensers trough the resistors. Each condenser represents one pixel in the image and the voltage is the pixel intensity (or color if we think of the voltage as an vector). If the condenser has lower voltage than its neighboring condenser, then it is being charged by that neighbor (the neighbor is being discharged by it) and vise versa. Charging speed depends on the resistance of the resistors. The greater the resistance is the slower is the charge/discharge rate.

Before we describe modification to the model, we have to understand the background itself and how it differs from foreground. Generally, the background consists of objects that are in most cases stationary and the foreground represents the *"moving"* objects. Basically, the background model represents the values occurring with higher probability and are not changing that much over time. The foreground objects are represented as values occurring less frequently and are significantly different from the background model. The background model can be statistically modeled using the Gaussian distribution. This method was presented by was proposed by Christopher Richard Wren et al. [3] and is known as temporal Gaussian background subtraction (TGBS). In this method, the mean value $\mu$ and standard deviation parameter $\sigma$ are modeled for each pixel from last $N$ images. Each new image is then compared with the model, and if new values are $2.5\sigma$ further away from the mean value, then it is marked as foreground. The model of the background could also be viewed as a time dependent signal filter, which is often realized as a combination of resistors and condensers in electrical engineering. This is also the basic idea behind our method.

### 3.1 Simple RC Model

Our method uses for background modeling a simulation of diffusion using a grid of condensers and resistors as proposed by by Pietro Perona and Jitendra Malik [4]. This model is modified in such a way, that each condenser representing one pixel of the background and the voltage over condenser is a representation of pixel intensity. This condenser is connected to excitation voltage representing the input image pixel values trough the additional resistor. The one-dimensional example of one block of the original filter and in comparison the modified filter can be seen in Figure 2. This way the video sequence values are filtered and the voltage $u_c$ over the condenser $C$ represents the filtered value (modeled background), which is an equivalent of the mean $\mu$ value used in TGBS. The standard deviation parameter $u_{c\sigma}$ can be modeled similarly using the absolute difference between input excitation voltages (values) and modeled background $u_c$ as an excitation voltage for $u_\sigma$ circuit. The differential equation describing the background update has the following form

$$\begin{aligned}\frac{\partial u_{c,x,y}}{\partial t} &= \frac{1}{CR}\left(u_{c,x-1,y} + u_{c,x+1,y} + u_{c,x,y-1} + u_{c,x,y+1} - 4u_{c,x,y}\right) \\ &\quad + \frac{1}{CR_e}\left(u_{e,x,y} - u_{c,x,y}\right) \ , \qquad (1)\end{aligned}$$

where the $u_c$ is the background model value (voltage over condenser), $x$ and $y$ represent the position in the circuit grid, $C$ is the condenser capacity, $R$ is the resistance value of the resistors connecting the condenser $C$ with the neighboring condensers, and the resistor $R_e$ connects the excitation voltage $u_e$ (new value) to the background model condensers. The magnitude of the excitation resistance

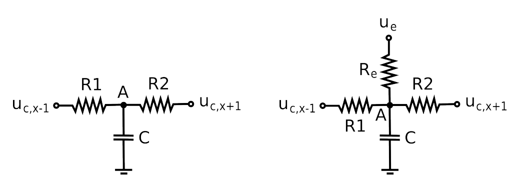



**Fig. 2.** The scheme of one block of the original (left) and modified filter (right)

$R_e$ regulates the speed of background model adaptation. If the resistance is too small, the adaptation speed is fast, and if it is big, the adaptation speed is slow. Fast adaptation leads to imprint of slowly moving objects into the background model. On the other hand, if the adaptation is too slow, some parts of the background image might be outdated.

This model is most similar to the TGBS method. The main difference between these two methods is in filtering. The TGBS filters the data only in time domain, because it filters the values for each input image pixel separately, but our method considers the background model as a whole, where all the pixels are connected to theirs nearest neighbors, which allows filtering also between background model pixels. The first experiment we propose compares the TGBS with our method. For the test, the level crossing dataset we have created is used. This dataset can be downloaded from our website `http://mrl.cs.vsb.cz/people/mozdren/levelcrossing/index.html` It consists of high resolution images and ground truth images for randomly selected frames. As a quantitative measure the Matthews Correlation (Phi) Coefficient (MCC) [8] is used. It computes the rate between true positive (true foreground), true negative (true background), False Positive (false foreground), and false negative (false background) pixels. The MCC is computed as follows

$$\Phi = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \ , \tag{2}$$

where $TP$ is true positive, $TN$ is true negative, $FP$ is false positive, and $FN$ is false negative. When the resulting coefficient is $+1$, the perfect prediction was measured, for $-1$ the inverse prediction, and for 0 the random prediction. The results can be seen in Table 1. As seen in the Table, our model performs better

**Table 1.** Performance comparison: RCBS: this model; TGBS: temporal gaussian

| | TP | TN | FP | FN | Phi |
|---|---|---|---|---|---|
| RCBS | 704612 | 19469974 | 80437 | 978641 | **0.59** |
| TGBS [3] | 630959 | 18605493 | 154090 | 1843122 | 0.41 |

than TGBS method. Further improvements for dynamic background adaptation, casted shadows removal, and segmentation filtering are presented in following parts of the text.

### 3.2 Layered RC Model

The simple layered RC model is already able to distinguish between foreground and background, but it still needs to be modified for adaptation to the dynamic backgrounds. Dynamic backgrounds are hard to adapt to. In many cases the background consists of moving objects like tree branches and water surface waves or objects frequently changing its color. The interiors, where the light are switched on and off are also considered as dynamic backgrounds. In this case, using only one Gaussian leads to high values of standard deviation $\sigma$. If the $\sigma$ is high, most of the input values are marked as background, even the moving objects in the foreground. This problem was solved by C. Stauffer and W.E.L. Grimson [7] in MoGBS. They developed a method, where the background is not modeled only by one Gaussian, but by a mixture of $K$ Gaussians. There, each of the Gaussians adapts to one kind of the backgrounds emerging in the video sequence.

Inspired by this method, we introduce the layered RC circuit model (LRCBS). We added to our circuit (model) additional layers, that are used to represent multiple backgrounds, the $DEMUX$ which is a demultiplex connecting the input voltages $u_e$ (input values) to specific layer, which is selected by selector $S$. Each layer is also connected by the inter-layer resistor $R_L$ providing inter-layer filtration. This helps mainly in initialization step, if the backgrounds are set randomly, and also when one background disappears. Layer, that does not represent any background moves towards to the nearest active background layer, where it helps to represent its background. The modified background modeler scheme can be seen in Figure 3. The adaptation of this model is driven by selector $S$. When

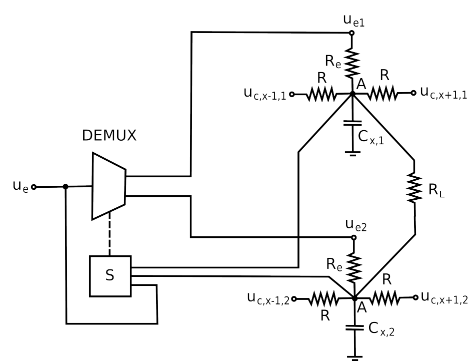



**Fig. 3.** The scheme of one one-dimensional block representing the layered RC model

the new excitation value $u_e$ emerges, the selector $S$ connects it using the demultiplex $DEMUX$ to the layer, where the absolute difference between new value $u_e$ and mean value $u_{c,l}$ is minimal. The other layers excitation values $u_{e,l}$ are leveled to theirs corresponding voltages $u_{c,l}$ (no excitation). This is performed for each block in the model and then the following difference equation is used for the update

$$u_{c,x,y,l}^{(t+1)} = u_{c,x,y,l}^{(t)} + \frac{dt}{RC}(u_{c,x-1,y,l}^{(t)} + u_{c,x+1,y,l,t}^{(t)} + u_{c,x,y-1,l}^{(t)}$$
$$+u_{c,x,y+1,l}^{(t)} - 4I_{c,x,y,l}^{(t)}) + \frac{dt}{R_e C}(u_{e,x,y,l}^{(t)} - u_{c,x,y,l}^{(t)})$$

$$+\frac{dt}{R_L C}(u^{(t)}_{c,x,y,l-1} + u^{(t)}_{c,x,y,l+1} - 2u^{(t)}_{c,x,y,l}) \ , \tag{3}$$

where $u^{(t)}_{c,x,y,l}$ is the value representing the background for pixels at position $x, y$ in the layer $l$ at time $t$. The $dt$ is a time difference constant, and $R_L$ is the resistance of the resistor connecting individual layers. Similarly, the $u_\sigma$ is modeled. The pixel is marked as foreground, if the input value is not within the distance of $2.5\sigma$ from the most similar layer value.

### 3.3 Selectivity

In some cases, the selectivity is introduced to background subtraction algorithms. It slows or stops the adaptation of the background for the input pixels marked as the foreground. This way, the moving objects do not imprint into the background that much or not at all. The selectivity is driven by the resistance of the resistor $R_e$ in our method. If the resistance $R_e$ is high, then the background adaptation is slow and vice versa. This implies that the resistance of the resistor should be driven by the difference between voltages $u_c$ and $u_e$. We have found the inspiration in a model of perceptron used in artificial neural network. Namely, the perceptron model used for back propagation neural network developed by David E. Rumelhart et al. [9]. The output of the perceptron $y$ is computed from the total input $x$ using the equation

$$y = \frac{1}{1 + e^{-\lambda(x-t)}} \ , \tag{4}$$

where $\lambda$ drives the slope of the function, and $t$ represents the threshold (the point, where the function moves rapidly from zero to one). Shape of the function can be seen in Figure 4. To use this function with our method, we have to

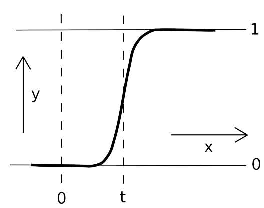



**Fig. 4.** The perceptron output function, used in back-propagation algorithm

define the maximal resistance $R_{max}$ (slow adaptation of the background) and minimal resistance $R_{min}$ (fast adaptation of the background). The function is then shifted up by $R_{min}$ and stretched by the difference between maximal and minimal resistances. The input value is given by the absolute difference (distance) between the excitation voltage $u_e$ and the condenser voltage $u_c$, and the threshold is given by modeled standard deviation $u_{c\sigma}$ multiplied by $T$, which is often set to 2.5 (represents 99 % of possible background values). This gives us a function with fluent transition between minimal resistance for values within the range $Tu_{c\sigma}$ and maximal resistance for values exceeding this range. This equation has the following form:

$$R_e = R_{min} + \frac{R_{max} - R_{min}}{1 + e^{-\lambda(\|u_e - u_c\| - Tu_{c\sigma})}} \ , \tag{5}$$

and the graph of the function can be seen in Figure 5.

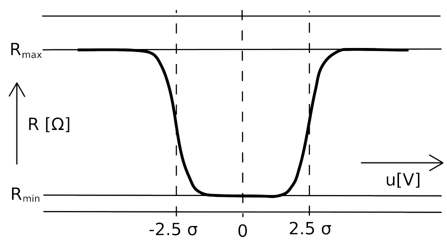



**Fig. 5.** Graph of a function for resistance regulation

### 3.4 Casted Shadows

Another problem is casted shadow. Objects moving in the video sequences often cast shadows and these are failingly marked as part of the object, which is not a wanted effect. In our method, the background/foreground segmented images are further processed by algorithm proposed by Thanarat Horprasert et al. [10]. They separate the brightness from chromaticity, and compute the brightness and chromaticity distortion using input pixel values and background model values. Those values are then thresholded and if those are within selected threshold, then the pixel is marked as shadow.

### 3.5 Segmentation Filtering

The background subtraction segmentations are often post processed by morphological operators. We have decided to use a more sophisticated filter. Our filter computes local histogram of segmentation affiliations for each pixel in segmentation and the affiliation for current pixel is substituted by the affiliation of the most frequent affiliation value (most probable value). The size of the area around the pixel directly affects the strength of the filter. The greater the area is, the stronger the filter is, and the more details are ignored. This filter can be used not only for binary segmentations, but also for ternary segmentations such as segmentation of moving objects, shadows and background. We have created an artificial example of segmentation to show ours filter abilities. The input and the output of the filter can be seen in Figure 6.

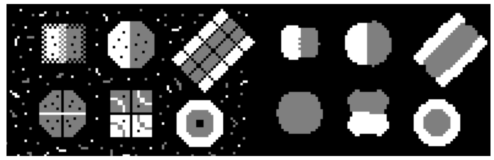


**Fig. 6.** Artificial segmentation before and after filtering

### 3.6 Frequency Sensitive Background Modeling

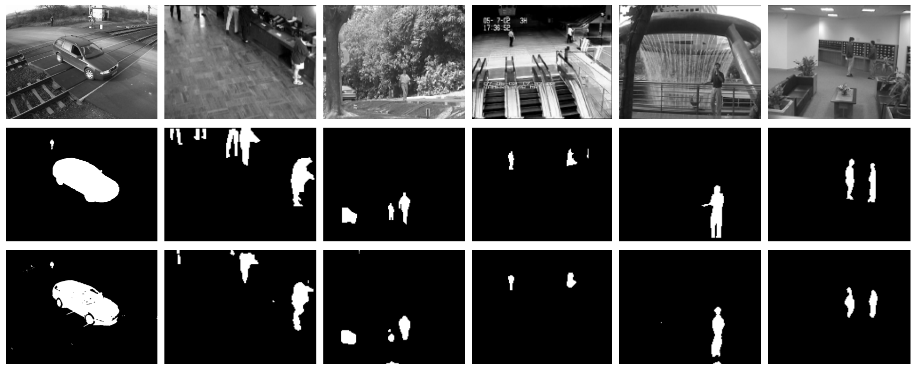
We have encountered one more problem. Our adaptation algorithm is similar to simple competitive learning [11], which in many cases becomes stuck in poor local solution (background). This also rises the problem, where some layers represent only small part of the background or none at all. We need each layer to represent approximately same quantity of the background. This is solved by introduction of a frequency sensitivity (frequency sensitive competitive learning) [12]. The basic idea is to store information about frequency of excitation of each layer in the block. In our method, we monitor the frequency of excitation, and if the pixel is marked as the foreground (new potential background), we excite the layer with the least frequency of the excitation. Consequently, this makes the least used layer to represent a new background.

## 4 Experiments

In this section we experiment with the complete method with all presented improvements and we compare it to other commonly used methods. The experiments were conducted on real video sequences captured from IP cameras installed on a level crossing, that we use for tests of obstacle detection and obstacle behavior monitoring. Furthermore, we use standard datasets, often used for testing and comparison of background subtraction algorithms. The level crossing dataset and ground truth images can be downloaded from `http://mrl.cs.vsb.cz/people/mozdren/levelcrossing/index.html`, and standard datasets with ground truth images created by L. Li [13] has been downloaded from `http://perception.i2r.a-star.edu.sg/bk_model/bk_index.html`.

The first experiment that we have conducted was a test of the abilities of the proposed method under different conditions like dynamic backgrounds and changing illumination. Resulting segmentations in Figure 7 clearly show that our method is able to adapt under many difficult conditions. The tests were conducted using the following configuration: $R = 2\ M\Omega$, $C = 1\ \mu F$, $dt = 0.001$, $R_L = 2\ M\Omega$, $R_{min} = 20\ k\Omega$, $R_{max} = 200\ k\Omega$, $\lambda = 5.0$, $K = 5$.

Our second experiment is quantitative. We compared our method with other commonly used methods. In this test, the level crossing dataset has been used. This dataset consists of high resolution images, and therefore, moving object detection can be measured more accurately. As a quantitative measure the previously presented Matthews Correlation (Phi) Coefficient (MCC) [8] is used. The



*a* *b* *c* *d* *e* *f*

**Fig. 7.** Testing images their foreground segmentations and ground truths. First row) input images; Second row) ground truth; Third row) Our segmentations, a) Level Crossing - exterior with shadows; b) Bootstrap - high movement, changing illumination; c) Campus - dynamic background, moving tree branches; d) Escalator - dynamic background; e) Fountain - dynamic background, water; f) Lobby - Strong change in illumination, switching lights on and off

results can be seen in Table 2. There, you can see that our method performs bet-

**Table 2.** Algorithms comparison: LRCBS: layered RC; MoGBS: mixture of gaussians; RCBS: Simple RC; TDBS: temporal difference; TGBS: temporal gaussian; TMBS: temporal median

| | TP | TN | FP | FN | Phi |
|---|---|---|---|---|---|
| LRCBS | 681225 | 20385616 | 103824 | 62999 | **0.89** |
| MoGBS [7] | 489913 | 20329902 | 295136 | 118713 | 0.7 |
| RCBS | 704612 | 19469974 | 80437 | 978641 | **0.59** |
| TDBS [5] | 461445 | 19816336 | 323604 | 632279 | 0.48 |
| TGBS [3] | 630959 | 18605493 | 154090 | 1843122 | 0.41 |
| TMBS [6] | 535682 | 17712377 | 249367 | 2736238 | 0.29 |

ter than other commonly used methods. The simple RC circuit model (RCBS) performs better, than TDBS, TGBS, TMBS, which are the methods, that are not able to adapt to dynamic background as well as RCBS. The layered version LRCBS is in addition able to deal with dynamic backgrounds. It is clear from the experiments that it outperforms the MoGBS.

## 5 Conclusion

We have developed a novel algorithm for background subtraction using a layered RC circuit for background modeling. We have shown, that our simple RC model performs better than TGBS method that is is similar to. This gave us good foundations for further improvements. The first improvement was introduction of additional layers, which allowed our method to represent dynamic backgrounds. We also dealt with casted shadows and output segmentation filtering. We have shown in the experiments that our algorithm is able to adapt to many difficult conditions like strong illumination changes, casting shadows, and dynamic backgrounds. Furthermore, our method performs better than other commonly used methods for background subtraction.

## Acknowledgement

This work was supported by the SGS in VSB Technical University of Ostrava, Czech Republic, under the grant No. SP2013/185, and Ministry of Industry and Trade of the Czech Republic - project TIP No. FR-TII/027 .

## References

1. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting (1995)
2. Papageorgiou, C., Poggio, T.: A trainable system for object detection. International Journal of Computer Vision **38** (2000) 15–33
3. Wren, C., Azarbayejani, A., Darrell, T., Pentland, A.: Pfinder: real-time tracking of the human body. Pattern Analysis and Machine Intelligence, IEEE Transactions on **19** (1997) 780 –785
4. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. Pattern Analysis and Machine Intelligence, IEEE Transactions on **12** (1990) 629 –639
5. Hironobu, A.L., Lipton, A.J., Fujiyoshi, H., Patil, R.S.: Moving target classification and tracking from real-time video. (1998) 8–14
6. Lo, B., Velastin, S.: Automatic congestion detection system for underground platforms. In: Intelligent Multimedia, Video and Speech Processing, 2001. Proceedings of 2001 International Symposium on. (2001) 158 –161
7. Stauffer, C., Grimson, W.: Adaptive background mixture models for real-time tracking. In: Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on. Volume 2. (1999) 2 vol. (xxiii+637+663)
8. Powers, D.M.W.: Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. Technical Report SIE-07-001, School of Informatics and Engineering, Flinders University, Adelaide, Australia (2007)
9. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. Cognitive modeling **1** (2002) 213
10. Horprasert, T., Harwood, D., Davis, L.S.: A statistical approach for real-time robust background subtraction and shadow detection. (1999) 1–19
11. Rumelhart, D.E., Zipser, D.: 5. In: Feature discovery by competitive learning. Volume 1. MIT Press, Cambridge, MA, USA (1986) 151–193
12. Ahalt, S.C., Krishnamurthy, A.K., Chen, P., Melton, D.E.: Competitive learning algorithms for vector quantization. Neural Networks **3** (1990) 277 – 290
13. Li, L., Huang, W., Gu, I.Y.H., Tian, Q.: Statistical modeling of complex backgrounds for foreground object detection. IEEE Transactions on Image Processing **13** (2004) 1459–1472