3D Extended Histogram of Oriented Gradients (3DHOG) for Classification of Road Users in Urban Scenes

Norbert Buch
norbert.buch@kingston.ac.uk
James Orwell
j.orwell@kingston.ac.uk
Sergio A. Velastin
sergio.velastin@kingston.ac.uk

Digital Imaging Research Centre
Kingston University
Kingston upon Thames, UK

8. Appendices

8.1. Logistics function for distance normalisation

A logistic sigmoid function is used to transform the distance measure d_k into a match measure m_k in the interval $[0,1]$. The function

$$m_k = \frac{1}{1+e^{a(b-d_k)}} \tag{1}$$

uses two parameters a and b for scale and shift of the transition region. The parameters can be estimated from the distance surface $\bar{\mathbf{D}}_{Mj}$ for every model. See Figure 1 for an example surface of an interest point. The values for this surface vary between interest points and therefore benefit from normalisation. The proposed parameterisation places the centre point of the distance surface d_C at the middle of the sigmoid function with a match measure of $m_k = 0.5$

$$\begin{aligned} m_k &= \left.\frac{1}{1+e^{a(b-d_k)}}\right|_{d_k=d_C} = \frac{1}{1+e^{a(b-d_c)}} \\ \frac{1}{2} &= \left.\frac{1}{1+e^{a(b-d_c)}}\right|_{b=d_c} = \frac{1}{1+e^{0}} \end{aligned} \tag{2}$$

which defines parameter $b = d_C$. The gradient of the sigmoid function at this point defines the other parameter a. The gradient should correspond to a line between the centre distance value at match measure $m_k = 0.5$ and the mean of all distance values $\bar{d}$ at match measure $m_k = 0$. This is illustrated as continuous blue line in Figure 2. The gradient of the line is

$$g = \frac{\bar{d} - d_C}{0.5} \tag{3}$$

which will be made equal to the gradient of the sigmoid function. The gradient of the logistics function at this point is

$$\left.\frac{dm_k}{dd_k}\right|_{\substack{d_k=d_C\\ b=d_C}} = \left.\frac{-1}{\left(1+e^{a(b-d_k)}\right)^2} e^{a(b-d_k)}(-a)\right|_{\substack{d_k=d_C\\ b=d_C}} = \frac{a}{2^2}, \tag{4}$$

which is made equal to the line gradient g :

$$\frac{a}{4} = g . \tag{5}$$

Using equation (3) and (5), the parameter a is given by

$$a = \frac{\bar{d} - d_C}{2} . \tag{6}$$

The resulting sigmoid function is illustrated in Figure 2. This function is used to convert distance measures $\mathbf{D}_{Mj}$ to match measures $\mathbf{M}_{Mj}$ during training and classification. An example output of the match measure $\mathbf{M}_{Mj}$ can be seen in Figure 3 showing a distinct peak, which is of the same height for all the interest points. By using the mean of all distance data points and therefore considering all data, a uniform drop of match measure is generated for different interest points. The impact of outliers is limited due to the bound output of the sigmoid function.

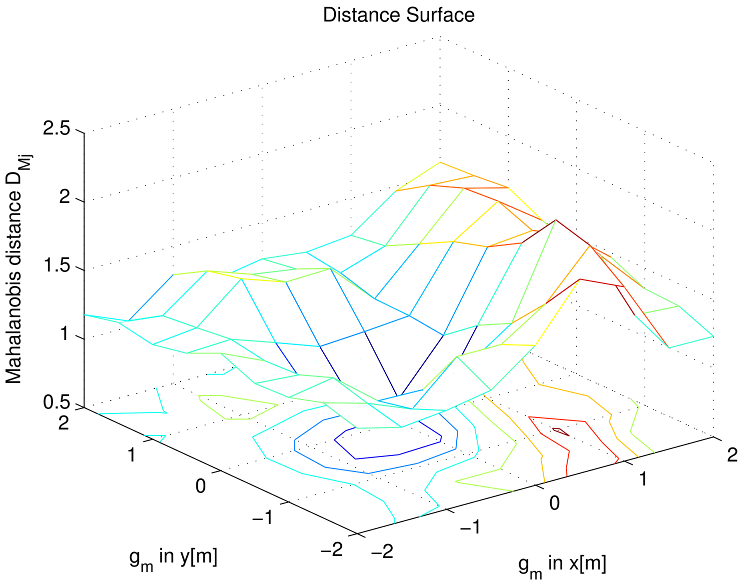

Figure 1 Example average feature distance surface. The centre at position $(0,0)$ corresponds to the training position and has usually the lowest value. The feature distance increases for coordinates further away from the centre.

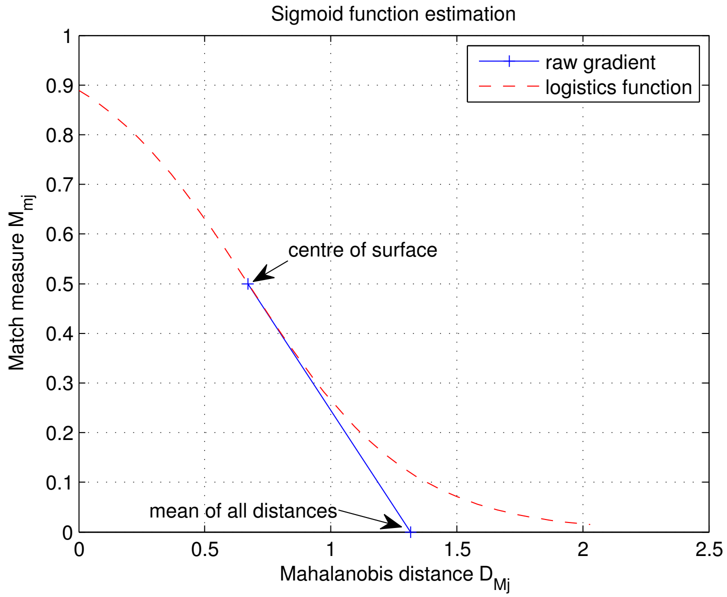

Figure 2 Estimated sigmoid function as dashed line. The continuous line is the gradient of the sigmoid function defined by the centre value of the distance surface and the mean distance of all grid points.

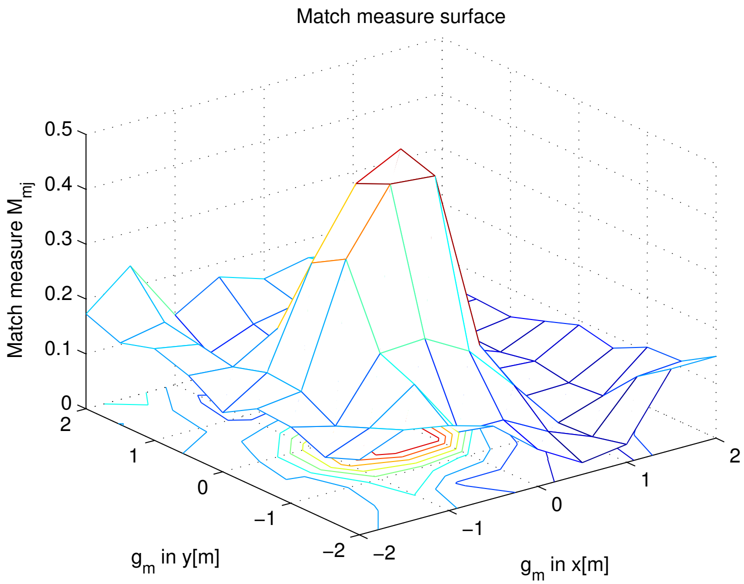

Figure 3 Final match measure surface after application of the sigmoid function. A distinct peak at the training position can be observed. This peak is set to the same value for all interest points.

8.2. Parameter influence

The influence of the 3D patch size is investigated. Reduced performance can be observed when reducing the patch to $\delta = 0.8\,\mathrm{m}$ at $\rho = 20\,\mathrm{P/m}$ and $\delta = 0.5\,\mathrm{m}$ at $\rho = 16\,\mathrm{P/m}$. Refer to Figure 4 for quantitative analysis.

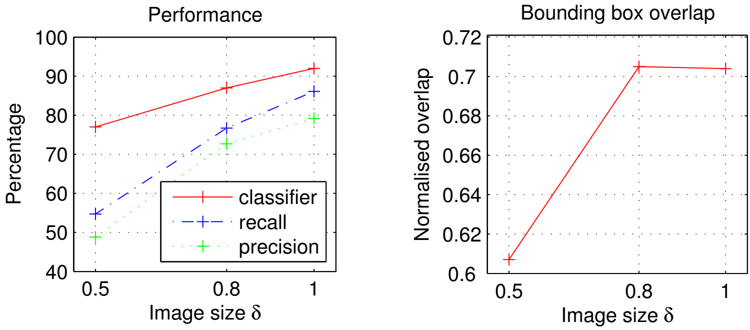

Figure 4 Performance increases in recall and precision for increased 3D patch size of 3DHOG classifier for oncoming driving direction.

8.3. Complete performance figures for three proposed features

This section provides all performance figures of the three algorithms proposed. An extended confusion matrix for full system performance, a confusion matrix of the classifier and a table with overall performance figures is given in Table 1 to Table 3.

ground truth / *detection*	bike	car/taxi	van	bus/lorry	FP
bike	1.00	.00	.00	.00	.44
car/taxi	.00	.83	.21	.03	.10
van	.00	.00	.67	.33	.08
bus/lorry	.00	.02	.02	.65	.00
FN	.00	.14	.10	.00	.00
count	27	361	48	40	
overlap	.70	.66	.73	.76	

ground truth / *detection*	bike	car/taxi	van	bus/lorry
bike	1.00	.00	.00	.00
car/taxi	.00	.97	.23	.03
van	.00	.00	.74	.33
bus/lorry	.00	.02	.02	.65
count	27	309	43	40

Symbol	*Value*
Recall *R*	81.1%
Precision *P*	82.0%
Classifier P_C	92.1%
Detector R_D	88.0%
Detector P_D	89.0%
GT Overlap	0.67

Table 1 Performance of 3DHOG feature. Left: full system performance; middle: classifier performance; right: cumulative figures.

ground truth / detection	bike	car/taxi	van	bus/lorry	FP
bike	1.00	.17	.34	.42	5.70
car/taxi	.00	.75	.39	.05	.32
van	.00	.02	.28	.07	.10
bus/lorry	.00	.00	.00	.47	.00
FN	.00	.06	.00	.00	.00
count	27	457	83	43	
overlap	.68	.54	.67	.80	

ground truth / detection	bike	car/taxi	van	bus/lorry
bike	1.00	.18	.34	.42
car/taxi	.00	.79	.39	.05
van	.00	.03	.28	.07
bus/lorry	.00	.00	.00	.47
count	27	430	83	43

Symbol	*Value*
Recall R	67.4%
Precision P	46.1%
Classifier P_C	70.5%
Detector R_D	95.6%
Detector P_D	65.4%
GT Overlap	0.57

Table 2 Performance of FFT feature. Left: full system performance; middle: classifier performance; right: cumulative figures.

ground truth / detection	bike	car/taxi	van	bus/lorry	FP
bike	.04	.04	.00	.03	.74
car/taxi	.04	.85	.61	.86	.31
van	.26	.04	.31	.00	.23
bus/lorry	.00	.00	.00	.11	.06
FN	.67	.08	.08	.00	.00
count	27	434	64	36	
overlap	.35	.58	.70	.67	

ground truth / detection	bike	car/taxi	van	bus/lorry
bike	.11	.04	.00	.03
car/taxi	.11	.92	.66	.86
van	.78	.04	.34	.00
bus/lorry	.00	.00	.00	.11
count	9	401	59	36

Symbol	*Value*
Recall R	69.9%
Precision P	57.9%
Classifier P_C	77.6%
Detector R_D	90.0%
Detector P_D	74.6%
GT Overlap	0.59

Table 3 Performance of histogram feature. Left: full system performance; middle: classifier performance; right: cumulative figures.