An Application of the Generalized Matrix Learning Vector Quantization Method for Cut-off-line Classification of Automobile-Headlights

Report 01/2015
Submitted: 10.01.2015
Published: 12.01.2015

S. Saralajew\textsuperscript{1,2}, Thomas Villmann\textsuperscript{2} and S. Söhner\textsuperscript{1}
\textsuperscript{(1)}Dr. Ing. h.c. F. Porsche AG, Weissach, Germany
\textsuperscript{(2)}University of Applied Sciences Mittweida, Fac. of Mathematics/Natural and Computer Sciences, Computational Intelligence Group, Mittweida, Technikumplatz 17, 09648 Mittweida, Germany

Machine Learning Reports
http://www.techfak.uni-bielefeld.de/~fschleif/mlr/mlr.html
Abstract

This paper deals with an application of the Generalized Matrix Learning Vector Quantization (GMLVQ) model for a calibration system of automobile-headlights. It describes the integration of GMLVQ into the Porsche Automatic Headlamp Setting (PAHS) system and reports first evaluation results about preliminary investigations. Thus, this paper is a good starting point for more detailed investigations in this field.
An Application of the Generalized Matrix Learning Vector Quantization Method for Cut-off-line Classification of Automobile-Headlights

S. Saralajew $^{1,2}$, T. Villmann$^2$, and S. Söhner$^1$

$^1$Dr. Ing. h.c. F. Porsche AG, Weissach, Germany
$^2$Computational Intelligence Group, University of Applied Sciences Mittweida, Germany

Abstract

This paper deals with an application of the Generalized Matrix Learning Vector Quantization (GMLVQ) model for a calibration system of automobile-headlights. It describes the integration of GMLVQ into the Porsche Automatic Headlamp Setting (PAHS) system and reports first evaluation results about preliminary investigations. Thus, this paper is a good starting point for more detailed investigations in this field.
1 Introduction and motivation

Trends in automobile-headlights show an increasing interest in intelligent lighting-systems. The high-performance LED-headlamps and the growing number of information, delivered by the specific sensors of the car, establish the base of these lighting-systems. The matrix-beam-system for example, which realizes a glare-free high-beam, is one of the first intelligent lighting-systems, which are available in current cars [Berlitz et al., 2014].

The performance of those systems depends especially on the interaction between the headlamp and the driver-assistance-camera. A correctly adjusted headlamp is as important as the calibration of the used camera. An online-calibration-system like PAHS\textsuperscript{1} [Söhner et al., 2013] can provide both demands. This system combines a movement of the headlamp’s light-distribution with a detection of the headlamp’s cut-off-line, see Fig.1, by the driver-assistance-camera. As a result of this calibration the horizontal and the vertical mis-aiming of the headlamp can be calculated. Thus, a continuous control of the headlamp setting is possible.

Due to this fact, a detailed consideration of the cut-off-line detection is demanded. Looking at [Flannagan et al., 1997] and [Manz, 2000] a change of the cut-off-line orientation can be observed. Several investigations have shown that the chromatic aberration of the lens generates a certain number $n$ of cut-off-lines, with a known fixed angle-offset in vertical and horizontal direction. Further, the final number $n$ depends on the headlamp and is fixed.

\textsuperscript{1}Porsche Automatic Headlamp Setting

![Figure 1: Cut-off-line of a headlamp.](image-url)
for each system. A detectable cut-off-line is marked by a sharp edge in a logarithmic gray-scaled luminance-measurement, see Fig.2. The angle-offsets between these cut-off-lines are fixed and do not depend on the environment conditions. For detailed information about the definition of the angle-offsets, the calculation procedure in PAHS, the modeling of a headlamp and the driver-assistance-camera see [Saralajew, 2014].

To expand the performance of PAHS it is important to have a mapping of the detected cut-off-line to the correct angle-offset.

2 Classification task

A detected cut-off-line in the PAHS system is considered as a data-vector $\mathbf{x}$ of six observable features $x_k$ of the cut-off-lines related to the canny-edge-detection and other environment conditions forming the data space $X \subset \mathbb{R}^m$ with $m = 6$. In particular, the components of data vector $\mathbf{x}$ are specified as follows:

- $x_1$ – threshold parameter of the canny-edge-detection;
- $x_2$ – minimal contrast-value of the scene during the detection of the cut-off-line;
- $x_3$ – maximal contrast-value of the scene during the detection of the cut-off-line;
- $x_4$ – cutting angle of the horizontal and vertical part of the cut-off-line.
- $x_5$ – length of the horizontal part of the cut-off-line;
- $x_6$ – length of the vertical part of the cut-off-line.

According to the scenario depicted in Fig.2, at least three different cut-off-lines types can be surely detected for the considered headlamp system:

- the upper cut-off-line labeled as $i = 3$;
- the lower cut-off-line assigned to the label $i = 2$;
- the so-called hotspot cut-off-line corresponding to the label $i = 1$.

Each of these cut-off-line types corresponds to a certain well-known fixed angle-offset. Because the calibration of PAHS is based on the initial detection of the lower cut-off-line, it is defined as zero-level whereas the other angle-offsets are in relation to this zero-level. In Fig.3 we depict an edge image obtained by the cut-off-line detection procedure identifying the upper and lower cut-off-line at the same time.

For the description as a mathematical classification system, the samples $\mathbf{x}$ of cut-off-line vectors form a data-space $X \subset \mathbb{R}^6$, which should be classified. For a predefined number $n$ of identified cut-off-line types we define the set of labels by $\mathcal{L} := \{1, \ldots, n\}$, whereby each label $i$ of $\mathcal{L}$ corresponds to a certain angle-offset as described above, i.e here we have $n = 3$.

Based on these conventions the classification task is to determine a classifier function

$$c : X \rightarrow \mathcal{L} : c(\mathbf{x}) \mapsto i,$$

such that a cut-off-line $\mathbf{x}$ is mapped onto a label $i$ and, hence, is assigned in this way to a certain angle-offset. According to the underlying physical measuring system generating the data vectors $\mathbf{x}$, this mapping depends on the luminance of the scene, the reflectance of the projection-surface as well as the distance between headlamp and surface.
3 Classification using the Generalized Matrix Learning Vector Quantization approach

The classification task (1) has to be determined by a machine learning tool. The tool used here is the Generalized Matrix Learning Vector Quantization (GMLVQ, see [Schneider et al., 2009, Kaden et al., 2014b]). The GMLVQ belongs to prototype based learning algorithms for classification, i.e. the algorithm adapt prototype vectors $w_k \in \mathbb{R}^n$ to training data according to the class distribution within the training samples. Mathematically, it is based on a dissimilarity evaluation

$$d_\Omega (x, w) = (x - w)^T \Omega^T \Omega (x - w)$$

(2)

between data and prototype vectors with a matrix $\Omega \in \mathbb{R}^{p \times n}$. This matrix $\Omega$ is denoted as scaling matrix or classification mapping matrix. The latter notation is due to the fact that the dissimilarity (2) can be rewritten as

$$d_\Omega (x, w) = (\Omega (x - w))^2$$

(3)

being the quadratic Euclidean distance of the linearly mapped vector difference $(x - w)$. When adapted for classification by GMLVQ training and $p \leq 3$, the mapping matrix offers an optimal classification visualization projection Bunte et al. [2012].
In summary, to determine a classification system based on GMLVQ we have first to generate a set of training data and subsequently to train the GMLVQ model.

### 3.1 Generating of the training data

The PAHS system offers the additional possibility to use a reference measuring-system [Söhner and Stork, 2014]. Based on the known distance between rear axle and projection surface it is possible to map the detected cut-off-line $x$ to the correct angle-offset. Using this angle-offset we can identify a desired class label $c(x)$ for each cut-off-line according to the pre-defined classes (cut-off-line types). In this way we are able to generate training data needed for GLMVQ training when generating data in a pre-determined setting.

To generate a valuable training data set the cut-off-line is detected in different well-defined characteristic scenes. During this process the data vector $x$ is recorded and the corresponding class $c(x)$ is defined by the reference measuring-system.

The generated set of training data consists of 1422 training vectors and presents the headlamp of Fig.1, i.e. $L$ is $\{1, 2, 3\}$. In the set of training-data there are approximately 13% of training vectors hot-spot cut-off-lines, 27% of upper cut-off-lines and 60% of lower cut-off-lines. A PCA-projection of the training-data is shown in Fig.4. The projecting eigenvectors are:

$$v_1 \approx (-0.4, -0.2, -0.9, 0, -0.1, 0)^T$$

$$v_2 \approx (0, 0, 0.1, -0.1, -1, 0)^T$$

Hence, the axes with highest data variance is a linear combination of “contrast-values” and the second significant axes is approximately the unit-vector of the length of the horizontal part. It is recognizable that the classes are not obviously separable in this two-dimensional projecting space. However, the different characteristic scenes during the recording process are reflected in the projection. For example, the right cluster of lower cut-off-lines is recorded using a black surface. Therefore the reflectance of the projection-surface is low and likewise the contrast-values. Additionally, the favorable
scene has a low luminance for the detection of the upper cut-off-line. Thus the upper cut-off-lines are placed near the lower cut-off-lines recorded on a black surface.

For the computing of the prototypes using GMLVQ the training-data is standardized with mean $\bar{x}_j$ and standard deviation $std(x_j)$. These values are given in Tab.1.

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$j$</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>$\bar{x}_j$</td>
<td>38.7665</td>
<td>13.2954</td>
<td>88.1301</td>
<td>19.2522</td>
<td>24.0850</td>
</tr>
<tr>
<td>$std(x_j)$</td>
<td>14.4787</td>
<td>10.1845</td>
<td>31.4416</td>
<td>4.9952</td>
<td>11.7099</td>
</tr>
</tbody>
</table>

Table 1: Standardization values for the training data.
3.2 Training of the GLMVQ model

The GMLVQ is an extension of Generalized Relevance Learning Vector Quantization (GRLVQ, [Bojer et al., 2001, Hammer and Villmann, 2002]) with origins in Learning Vector Quantization models proposed by Kohonen in [Kohonen, 1988, 1995]. Training of the GMLVQ is a stochastic process triggered by random presentation of training data. Assuming data vectors \( x \in \mathbb{R}^m \) with class labels \( c(x) \in \mathcal{L} \), GMLVQ requires the set \( W = \{ w_k \in \mathbb{R}^m, k = 1 \ldots M \} \) of prototype vectors with class label \( y_{w_k} \in \mathcal{L} \), at least one per class. The cost function minimized by GMLVQ is

\[
E_{GMLVQ}(W) = \frac{1}{2} \sum_{v \in V} f(\mu(v))
\]

approximating the classification error [Kaden et al., 2015], where

\[
\mu(x) = \frac{d_{\Omega}^+(x) - d_{\Omega}^-(x)}{d_{\Omega}^+(x) + d_{\Omega}^-(x)}
\]

is the classifier function. We remark that \( \mu(x) \in [-1, 1] \). Further \( d_{\Omega}^+(x) = d_{\Omega}(x, w^+) \) denotes the dissimilarity between the data vector \( x \) and the closest prototype \( w^+ \) with the same class label \( y_{w^+} = c(x) \), and \( d_{\Omega}^-(x) = d_{\Omega}(x, w^-) \) is the dissimilarity degree for the best matching prototype \( w^- \) with a class label \( y_{w^-} \) different from \( c(x) \). The transfer or squashing function \( f \) is a monotonically increasing function usually chosen as sigmoid or the identity function [Sato and Yamada, 1996].

Learning in GLVQ is performed by the stochastic gradient descent learning for the cost function \( E_{GLVQ} \) with respect to the winning prototypes \( w^+ \) and \( w^- \) for a randomly chosen training sample \( x \in X \) and randomly initialized prototypes. The updates can be written as

\[
\Delta w^\pm \propto -\frac{\partial f}{\partial \mu(x)} \cdot \frac{\partial \mu(x)}{\partial d_{\Omega}^\pm(x)} \cdot \frac{\partial d_{\Omega}^\pm(x)}{\partial w^\pm}.
\]

The prototype update is accompanied by the adaptation of the scaling matrix \( \Omega \) of the dissimilarity measure \( d_{\Omega} \) from (2) to achieve a better classification performance. Starting with the unit matrix \( \Omega = I \) it is subsequently
adapted according to
\[ \Delta \Omega_{ij} \propto -\frac{\partial f}{\partial \mu(x)} \cdot \left( \frac{\partial \mu(x)}{\partial d^+} \cdot \frac{\partial d^+}{\partial \Omega_{ij}} + \frac{\partial \mu(x)}{\partial d^-} \cdot \frac{\partial d^-}{\partial \Omega_{ij}} \right) \] (7)

at the same time as the prototype update takes place Schneider et al. [2010].

In the classification problem to be considered here, we used \( M = 9 \) prototypes, three for each class in ascending order. The achieved prototypes are

\[
\begin{align*}
\mathbf{w}_1 &= (1.1975, 1.4491, 0.9626, -0.0095, 0.3155, -0.1545)^T, \\
\mathbf{w}_2 &= (0.6836, -0.0660, 0.9007, -0.1051, -0.0360, -0.0402)^T, \\
\mathbf{w}_3 &= (-1.1815, -0.6942, -1.4685, -0.1642, -0.6036, -0.1494)^T, \\
\mathbf{w}_4 &= (-1.0377, -1.0158, -0.7847, -0.5099, -0.4270, -0.7007)^T, \\
\mathbf{w}_5 &= (0.4686, -0.9196, 1.1873, 0.6263, 0.3441, -0.5295)^T, \\
\mathbf{w}_6 &= (-0.9950, -1.0253, -0.7424, 0.3656, -0.1556, 0.9506)^T, \\
\mathbf{w}_7 &= (1.3887, 0.9023, 1.3267, 0.5165, 0.0003, -0.7472)^T, \\
\mathbf{w}_8 &= (1.2769, 0.8930, 1.3262, -0.1375, -0.0756, 1.1646)^T, \\
\mathbf{w}_9 &= (1.3878, 0.8330, 1.7093, -1.1449, 0.9115, 0.4148)^T.
\end{align*}
\]

The mapping \( \mathbf{\Omega} \)-matrix considered in (3) was obtained as

\[
\begin{pmatrix}
0.8220 & -0.0875 & -0.0550 & -0.0758 & 0.0147 & -0.1545 \\
-0.0641 & 1.7095 & -0.4775 & -0.0556 & 0.0382 & -0.0402 \\
-0.0916 & -0.5636 & 1.3072 & -0.0574 & 0.0413 & -0.1494 \\
-0.0616 & -0.1495 & 0.0344 & 0.1427 & -0.0332 & -0.7007 \\
0.0096 & -0.0428 & 0.0851 & -0.0211 & 0.0668 & -0.5295 \\
-0.0509 & -0.0739 & -0.0241 & 0.0743 & 0.0046 & 0.9506
\end{pmatrix}
\] (8)

realizing the linear mapping to achieve best classification performance. Finally, the GMLVQ-classifier achieved an overall accuracy performance of 97.82%, which indicates an approximately perfect class separability. Without metric adaptation the accuracy was only in the range of about 80% reflecting the importance of metric adaptation.
3.3 Application in a real system for online calibration

After training of the GMLVQ-classification model it can be used in the PAHS system for cut-off-line classification in real situations. If a new cut-off-line $\mathbf{x}$ is detected by the PAHS system, the classification procedure requires three calculation steps based on the obtained GMLVQ-prototypes, see (3.2), the GMLVQ-computed mapping matrix $\Omega$, see (8), and the standardization values from Tab.1:

1. standardization of $\mathbf{x}$:
   
   $$x_j := \frac{x_j - \bar{x}_j}{\text{std}(x_j)} \text{, for } j = 1, 2, \ldots, 6$$

2. transformation of $\mathbf{x}$ using the scaling matrix $\Omega$:
   
   $$\mathbf{x} := \Omega \cdot \mathbf{x}$$

3. calculation of the best-matching prototype $\mathbf{w}_i$ using the distance $d_{\Omega}$ and assign the label $y_{\mathbf{w}_i}$ to the class label $c(\mathbf{x})$, i.e.
   
   $$c(\mathbf{x}) := y_{\mathbf{w}_i^*} \text{, with } i^* = \text{argmin}_{i=1,\ldots,M}(d_{\Omega}(\mathbf{w}_i, \mathbf{x}))$$

If the class label $c(\mathbf{x})$ is calculated, then there are two possibilities. The first is that a detected cut-off-line $\mathbf{x}$ with a class unequal to the desired class is ignored. This means that the detected cut-off-line $\mathbf{x}$ is not considered in PAHS. The disadvantage of this method is that the number of possible cut-off-lines for PAHS is decreased. But the advantage is that angle-offsets are not needed. The second method is to correct the detected cut-off-line by using the corresponding angle-offset. Thus, the benefit is that all detected cut-off-lines are usable in the online-calibration. However, the angle-offsets are needed and an error in the angle-offsets has a direct influence to PAHS.

In a real application of the online-calibration the probability that $\mathbf{x}$ corresponds to $c(\mathbf{x})$ is like the probability in section 3.1. Therefore the ignoring of detected cut-off-lines unequal $c(\mathbf{x}) = 2$ is not acceptable. The application
of the correction step using the GMLVQ and the angle-offsets show a huge benefit in the application.

Clearly, the efficiency of the classification depends in a highly manner on the quality of the training data. The used training data consist of measurements on

- a black surface under dark environment conditions.
- a white surface under dark and light environment conditions.
- a gray surface under light environment conditions.
- a red surface under light environment conditions.

In preliminary experimental studies performed so far, the GMLVQ model is also able to classify those cut-off-line of scenes correctly, which are not included in the training data. However, not all scenes were classified correctly. For this purpose, rejection strategies as proposed in [Fischer et al., 2014] should be included in GMLVQ model in the future.

4 Outlook

The application of the GMLVQ for classifying the cut-off-line shows a high potential. However, before an application can operate in a production vehicle, further investigations need to be done. Due to the fact that for each headlamp-system the determination of the GMLVQ prototypes and the scaling/mapping matrix $\Omega$ have to be done just one time, it is necessary to define uniform measurement conditions for the generation of the training data. Moreover it is necessary to perform further studies regarding the difficulty of scene selection important to build up a classifier with high generalization ability for the classification. The goal is that both the set of training data and the classification error in the real world application is as small as possible. Further, as previously mentioned, appropriate rejection strategies have to be implemented.
A huge advantage of the GMLVQ (and also of other related LVQ methods) is that the learning (adaptation) of the prototypes is done in off-line mode. Hence, the time processing costs are not important in the application mode. The calculation costs for the classification with pre-determined prototypes are negligible in comparison to the general calculation costs of the PAHS calibration. Therefore, the described system is principally applicable in a production vehicle, from this point of view.

The presented procedure is based on all six chosen features of the cut-off-line so far, see Sec.2. Yet, GMLVQ provides further information. The matrix

$$\Lambda = \Omega^T \Omega$$

is the so-called classification correlation matrix. This matrix offers the additional information, which data dimensions and combinations thereof contribute to a good classification performance. For the considered cut-off-line classification system it is displayed in Fig.5. The appraisal of this matrix
could be used to decrease the dimension \( n \) of the cut-off-line vectors for further decrease of computation time. Otherwise, this the information contained in the classification correlation matrix could be used to define the position of the contrast-values for the cut-off-line, such that they are at the most significant positions for a successful classification.

Further studies could include also GMLVQ-variants optimizing other statistical measure than classification error [Kaden et al., 2014a] or optimize the receiver operating characteristics (ROC), when only two types (classes) of cut-off-lines are to be classified [Biehl et al., 2014].
References


M. Kaden, W. Hermann, and T. Villmann. Optimization of general statistical accuracy measures for classification based on learning vector quantization. In M. Verleysen, editor, Proc. of European Symposium on Arti-


