https://doi.org/10.2352/ISSN.2470-1173.2018.17.AVM-147

Visual quality evaluation of the multi-camera visualization in automotive surround view systems

Vladimir Zlokolica, Mark Griffin, Aidan Casey, Daniela Solera, Brian Deegan, Patrick Denny, Barry Dever; Valeo Vision Systems, Tuam, Galway, Ireland; University of Novi Sad, Serbia

Abstract

Surround view camera systems are nowadays commonly provided/offered by most of the car manufactures. Currently, a considerable number of different multi-camera visualization systems exist in the automotive sector, which are difficult to evaluate and compare in terms of visual performance. This is mainly due to the lack of standardized approaches for evaluation, un-predictable 3D input content, un-predictable outdoor conditions, non-standardized display units as well as visual quality requirements that are not clearly identified by the car manufactures. Recently there has been IEEE-P2020 initiative established that concerns standards for image quality for automotive systems. In this paper, we address the problem of reliably evaluating multi-camera automotive surround view systems in terms of visual quality. We propose a test methodology and an efficient test system platform with a video playback system and real camera input images captured from the vehicle, which enables visual quality monitoring subjectively on the head unit display and objectively by the proposed objective quality metrics.

Introduction

Imaging in automotive systems [1] has become an important part of the whole automotive system, providing the ability of viewing and perceiving the 3D scene surrounding the vehicle. Advanced automotive imaging systems provide smart driver-assistance systems consisting of multiple cameras that can perform complex 3D visualization [2, 3], sensing and recognition tasks [4] that are used for advanced security, driver road assistance and autonomous driving.

Testing and performance evaluation of such systems is difficult due to various factors such as uncontrollable outdoor environments and the complexity of the outdoor scene. As such, imaging in the automotive environment is particularly demanding and needs to be first evaluated [5, 6] at the camera component level. Although there are some existing standards for evaluating camera imaging systems, such as ISO 15739 (for noise measurement), ISO 12233 (for sharpness measurement), 1858 - Standard for Camera Phone Image Quality (CPIQ) and EMVA 1288 [6] they still need to be adapted to the specific automotive environment. One way to verify this is in a lab environment by modelling the photographic space [7] of the automotive environment conditions, which can then be applied to simulate real-life situation.

The second step in visual performance evaluation of the automotive vision system is to evaluate the complete multi-camera system with the embedded system processing unit, usually referred to as the ECU in automotive sector. This assumes performance validation of the whole processing chain going from the multiple camera devices to the output video to be displayed on the head unit display, or to the output of specific computer vision task. In such framework, each of the functionalities of the automotive vision systems have to be designed independently to a certain extent, since each of the tasks: (i) visualization [3], (ii) computer vision tasks [8, 1], (iii) security warnings [9] and (iv) autonomous driving have different requirements.

In this paper, we are focusing on the multi-camera visualization aspect for driver assistance, where special attention is paid to the visual quality of the multi-camera views displayed on the vehicles' display. Such visual quality performance evaluation of the advanced automotive vision systems is challenging because of the missing standardized approaches for evaluation, un-predictable input 3D content, un-predictable outdoor conditions as well as non-standardized display units. In our work, we optimize the input multi-camera image quality to a sufficiently high level and use that as an input to the ECU processing device, which outputs the optimized multi-camera surround-view video to the head unit display in the vehicle. The output video captured is analysed in terms of objective quality metrics, which are then correlated to the subjective opinion scores to model optimally true driver visual perception.

The main objective in this paper is to propose a reliable test system framework and methodology for visual quality evaluation of multi-camera surround view automotive systems. The second objective is to determine adequate test case scenarios in terms of the 3D content and environment conditions, which can be representative for surround view systems visual quality evaluation. The final objective is to develop reliable objective metrics to efficiently monitor visual quality performance, in terms of the defined Key Performance Indicators (KPIs) on the proposed test system framework.

This paper is organized as follows. We first explain the proposed test system platform and general methodology for the visual quality performance evaluation. After that we discuss and propose a 3D content and environment conditions selection method for the input multi-camera captures, to be used as an input to the test system. Finally, we propose visual quality metrics for the surround view images for analysing output video. We conclude the paper by providing experimental results.

The proposed test system platform and methodology

We propose a test system framework and methodology for visual quality performance of the surround-view systems. This includes the design of the controllable input data module and video capturing module for output video from the surround view system under evaluation. The input multi-camera module is used to stream previously recorded input multi-camera videos with all

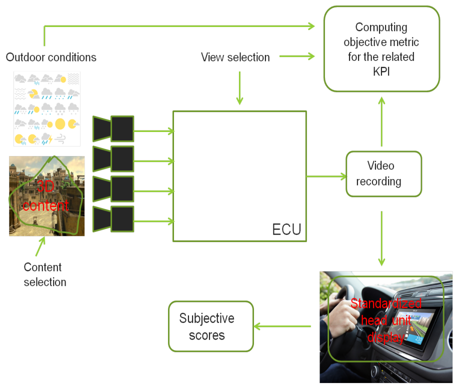

Figure 1. *Test system framework overview.*

supporting vehicle data to an ECU. The input multi-camera images/videos are previously recorded in the vehicle for selected outdoor scene and outdoor conditions.

The output video capturing module enables storing the output of the multi-camera surround view under inspection. Such surround view output video can then be streamed to the objective metric computation module, where one can monitor the status of the ECU output in terms of the surround view visual quality. Additionally, the captured output surround view video is streamed simultaneously to the head unit display for subjective visual quality assessment. The proposed test system framework is shown in Fig. 1.

In the lab, subjective visual quality assessment is carried out on a specific display unit, which represents a standard monitor for all ECU surround view systems evaluation. The viewing conditions are designed in such a manner to closely match the viewing conditions of the driver in the vehicle.

Both objective and subjective quality assessment can be done in real-time or off line for debugging purposes in cases of unexpected output. The captured surround view video sequence is analysed in terms of different attributes, such as brightness and colour harmonization, noise level, sharpness. In addition, for the ECU testing, a double viewport screen was developed to render an uncorrected merged view next to a merged view with photometric corrections applied. Such double surround view images/videos can then be used for easier subjective and objective quality evaluation. In the case of subjective visual quality evaluation on the display screen, the testing can then be performed following already existing standards such as [10].

Furthermore, the proposed evaluation methodology allows for efficient testing of different types of surround views (e.g., TopView or BowlView), using the same input multi-camera sequences. For different types of views we model subjective perception and adapt the quality metrics to the subjective visual quality assessment. This particularly relates to objective quality metrics, which have to be adapted to a different subjective perception in different types of the surround views. Specifically, for different surround view types, various quality metrics and KPIs are investigated and analysed in terms of their correlation and relevance to the final overall visual quality.

Finally, we also take into account the 3D content type and environment conditions under which the multi-camera system is exposed to, to determine overall visual quality assessment. This is considered as a type of support for deciding which of the quality metrics are most relevant in particular cases and how they should be optimally combined to obtain reliable overall visual quality estimates.

3D scene content and automotive environment condition selection

We analyse different scene case scenarios that the input multi-camera system is exposed to, which can be used as a selected set of inputs to the surround-view system under evaluation. In a specific environment there are two important features to be considered when selecting data sets for testing: 1) 3D content to be captured and 2) environment conditions under which the content is acquired by multi-camera system.

For the scene content selection, one approach is to select relatively constant environment conditions and capture different scenes, after which the methodology of [10] can be applied, for example. Once the 3D content is well selected, we record each selected scene under different lighting/environment conditions. The proposed scheme for capturing the same scene under different environment conditions is based on the photographic space concept, initially proposed by Eastman Kodak [5, 7]. In its simplest form, photographic space can be described as a two-dimensional map [7], where one dimension is camera-to-subject distance d and the other dimension is the luminance L or brightness of a scene.

Photographic space information can be used as an input to specifying under which conditions the visual quality assessment has been performed and as such be reproducible to a certain extent. The authors in [7] introduce automotive photographic-space definition to be used for specifying different lighting conditions under which the camera is exposed to. In this paper, we aim at extending this concept to multi-view camera systems, where each camera can be in different photographic space position under the same outdoor conditions.

Using the proposed approach we select a standard scene catalogue that we use for capturing input multi-camera videos, which are then subsequently used for visual quality evaluation of the target surround view systems, with different types of the views, i.e., different view use-cases.

Visual quality metrics for the surround view

In this section we discuss visual quality assessment of the surround view images in terms of brightness and colour harmonization; noise and sharpness assessment metrics goes out of the scope of this conference paper. We propose novel approaches for brightness and colour harmonization objective visual quality metric.

Metrics for brightness balancing and colour harmonization visual quality assessment

In this subsection we propose novel approaches for visual quality assessment of surround view images in terms of brightness balancing and colour harmonization. As an input we use both harmonized and non-harmonized surround view image textures. We describe the proposed approach using the TopView of a 4 camera surround view system; we note that extension to other surround view images would be straightforward considering different regions of interest (ROI).

As a first step in our approach we divide TopView images

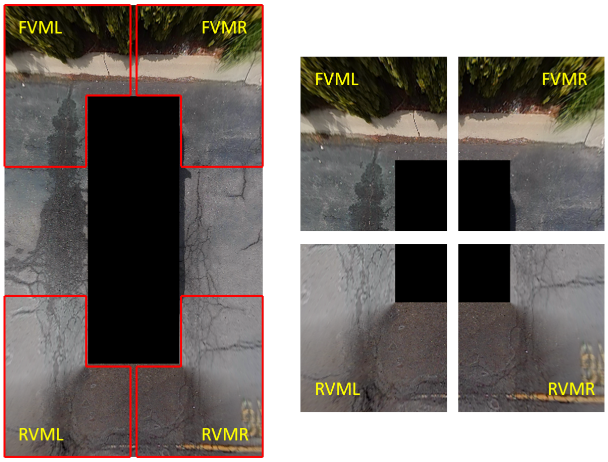

Figure 2. *Dividing TopView image into four corner image textures for the non-harmonized output; FVML, FVMR, RVML and RVMR correspond to Front-Left, Front-Right, Rear-Left and Rear-Right merged corner camera images respectively.*

(harmonized and non-harmonized) into four parts corresponding to textures in 4 corners, as shown in Fig. 2, for the un-harmonized case. As can be seen in the shown images, we define four corners: FVML, FVMR, RVML and RVMR, corresponding to Front-Left, Front-Right, Rear-Left and Rear-Right merged camera images respectively. We subsequently analyse those four corners of the surround view images in terms of photometric distribution in order to determine metrics for brightness and chroma correction. We perform analysis in YUV space, where the Y component is used for brightness harmonization assessment and the UV chroma components are used for colour harmonization assessment.

Metric for brightness balancing visual quality assessment

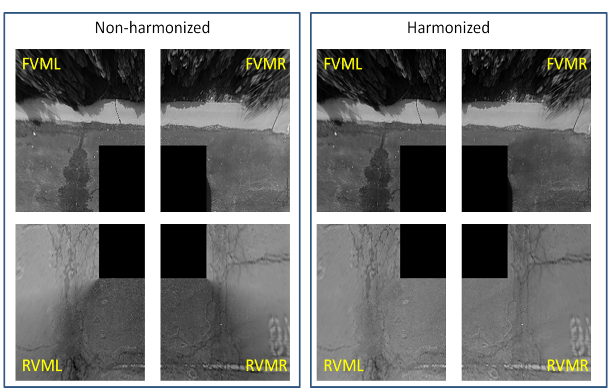

Figure 3. *TopView luminance image textures of the four corners for the harmonized and non-harmonized output.*

The metric for brightness balancing quality assessment is computed using the described four corner image textures, using the luminance only component. We use luminance textures from both harmonized and non-harmonized image textures (Fig. 3) and compare those in terms of the proposed metric, to determine if the harmonized image has improved the visual quality or not, as well as to which extent. Besides the relative difference between the harmonized case and non-harmonized one, we also propose a threshold which indicates that the harmonization quality is acceptable or not.

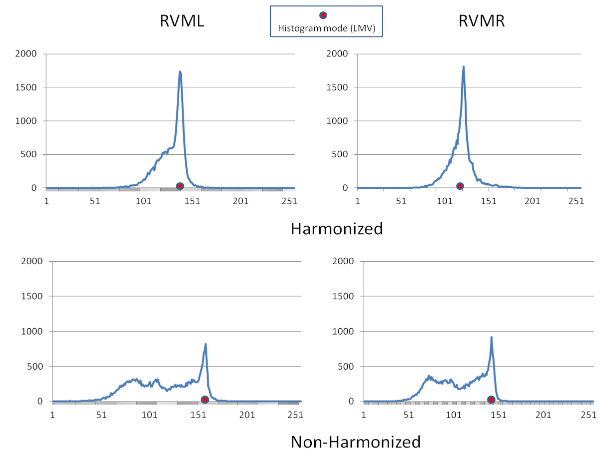

Figure 4. *Histograms of the RVML and RVMR luminance corner textures corresponding to Fig. 3 for the harmonized and non-harmonized case.*

We compute histograms for all the corner images and analyse them to compute the proposed deviation metric, which indicates the brightness harmonization quality. The histograms of the harmonized and un-harmonized textures for the RVML and RVMR corners, for the textures shown in Fig. 3, are given in Fig. 4. As can be seen from the histograms in Fig. 4, the histograms for the non-harmonized view textures are more spread across different luminance values of the RVML and RVMR textures (Fig. 3), in comparison to the harmonized case. As such, we can model the luminance discrepancies through histogram distribution - it can be seen that in the harmonized case larger portion of the luminance values are closer to each other, i.e., close to the certain peak values, which represents most dominant part of the particular region.

Considering this, in the proposed brightness balancing quality methodology we first compute the peak position, i.e., histogram mode and define it as the Luminance Mode index Value (LMV); the LMV of the histograms are shown with red dots in Fig. 4. We assume the LMV is representative of the luminance of the area of interest within the scene. Preferably, such computed LMV should be corresponding to the road surface, which is considered to be the most common reference for computation in automotive systems; nevertheless, it can also be corresponding to other content such as grass surrounding the vehicle as long it represents the most dominant content. Further on, we introduce a constraint on computing LMV, by restricting the LMV to be between low histogram index (LHI) and high histogram index (HHI) values. Such a constraint is used assuming that representative values should be in some mid range, e.g., LHI = 50, HHI = 200; we note however that precise values are determined based on experimental results.

After we determine the LMV, we compute the histogram deviation around the LMV value within the specifically computed region of interest, in the histogram domain. More specifically, we compute the deviation, using the LMV as a reference point and consider only histogram values to the left and to right side, for the assigned Δ value, i.e., correspondingly computed Δ_L and Δ_H

values, as follows:

$$HD = \sqrt{\frac{\sum_{i=LI}^{HI}(hist(i)(i-LMV)^2)}{\sum_{i=\Delta_L}^{\Delta_H} hist(i)}} \qquad (1)$$

where HD determines histogram deviation and Δ value is determined experimentally and ranges roughly from 50 to 100. Additionally, $LI = LMV - \Delta$ and $HI = LMV + \Delta$, where LI and HI are bounded, i.e., clamped to be within the $[LHI, HHI]$ range.

The histogram deviation (HD) is computed for all corner luminance image textures and corresponding histograms (FVML, FVMR, RVML, RVMR), separately for the harmonized and unharmonized case. The final estimated brightness balancing metric (BBM) value for the whole TopView image is then computed as a weighted sum of the histogram deviation values for all four corner images, where weighting is determined through a confidence score that the corner brightness harmonization assessment value can be estimated reliably enough.

Metric for colour harmonization visual quality assessment

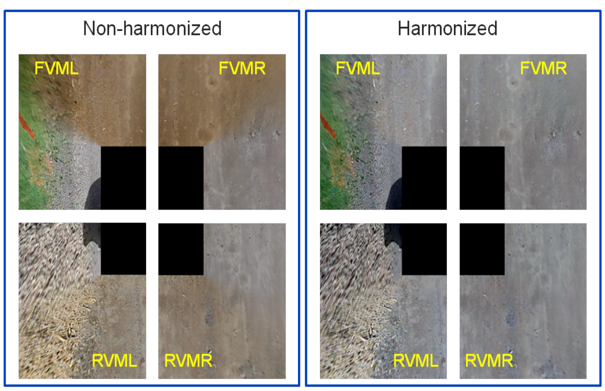

Figure 5. *Corner Textures from both non-harmonized (left) and harmonized (right) TopViews.*

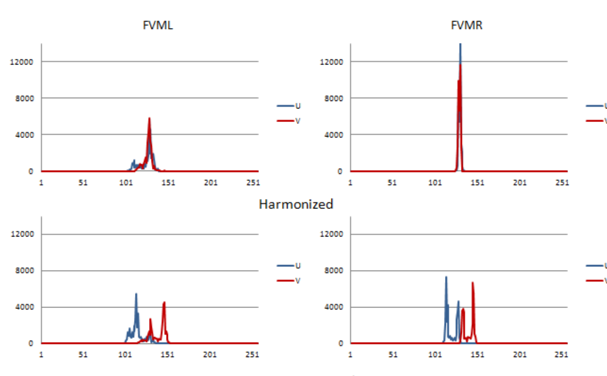

Figure 6. *Histograms of U and V image components for the non-harmonized and harmonized FVML and FVMR corner textures*

The metric for colour harmonization quality assessment is computed using the corner image textures, from the surround view images, as shown in example in Fig. 2. In the proposed metric for colour correction, we work in UV space within $U,V \in [0-255]$ and where neutral "grey" chroma values corresponds to $U,V = 128$. We extract the U and V chroma components from the corner textures and process them to determine the degree of colour hue and colour dis-harmonization between different camera texture images, within the surround view.

For each corner image we first compute the U and V histograms and analyse them to determine the degree of harmonization/un-harmonization. In Fig. 5, we show another example of corner textures, obtained in the same manner as depicted in Fig. 2, which contains high degree of colour hue and colour dis-harmonization. We present this example for easier analysis and explanation of the difference between the un-harmonized and harmonized TopView images, in terms of the colour cast. The corresponding histograms of the U and V components for the harmonized and non-harmonized corner textures FVML and FVMR are additionally shown in Fig. 6.

The U and V histograms (Fig. 6) for the FVML unharmonized corner texture (Fig. 5) clearly display the orange cast seen in this texture. The histogram peaks are below 128 for U and above 128 for V indicating the orange cast from the front camera input, the rest of the histogram represents the grey road and grass patch from the ML camera input. The histograms of the harmonized image have both U and V merged back to around 128 with the small peak of U below 128 representing the grass patch. This indicates the FVML corner texture has been harmonized to a certain degree. The FVMR corner texture has no other colour textures in the image so the histograms only represent the road. The peak of U below 128 and the peak of V above 128 representing the FV camera input. The second peaks for both U and V are much closer to 128 representing the MR camera input. The histograms for the harmonized FVMR corner texture for U and V have shifted back towards 128.

Considering this brief analysis, we aim at developing a metric that will measure the degree of the colour harmonization distortion (dis-harmonization), i.e., to what degree is the surround view harmonized and to what degree is the colour cast removed, in comparison to the un-harmonized case. In the proposed approach we compute the mathematical expectation of U and V values, considering only the histogram bins and corresponding histogram values around the reference grey value ($U,V = 128$); we use the histogram data that is within the range of +/- a given Δ from the reference grey value. We make the assumption that the road is always grey or close to grey, therefore any data points of the histogram outside of this region will be deemed as colour objects or colour textures in the vehicles environment that we do not want to include in our analysis. We then calculate the euclidean distance from the centre (128,128) of the UV colour-space to the U and V mean point for each corner texture. This euclidean distance we can term as our Local Colour Mean (LCM) and provides information about mean local hue of the image corner texture, after applied harmonization.

We then calculate the variance (second moment) of the data within a determined range, relative to the estimated LCM value. The estimated variance value, we refer to it further as Local Colour Variance (LCV), is expected to provide information about how well the corner texture is harmonized and describe certain confidence score as to what degree the colour hue is removed by applied harmonization. Specifically, the LCV describes the vari-

Table 1: Results of the brightness and colour correction harmonization evaluation: H - harmonized surround view, UH - un-harmonized surround view, LF - low spatial frequency, MF - medium spatial frequency, HF - high spatial frequency, LDR - low dynamic range, MDR - medium dynamic range, HDR - high dynamic range, BBHD - high level of brightness harmonization distortion, CCHD - high level of colour harmonization distortion. The BBMetric and CCMetric values represented in bold correspond to cases shown in Fig. 3 and Fig. 5, respectively.

Seq.	Sequence description	Conditions	BBMetric		CCMetric	
			UH	H	UH	H
1	City road with shade LF-MDR	Partly sunny, left image very bright BBHD and CCMD	29	18	5.7	1.3
2	Parking and trees MF-MDR	Overcast, rear image darker than other 3 BBHD and CCMD	**37**	**20**	4.1	1.6
3	Parking between cars HF-MDR	Sunny, rear and front image darker than other 3 BBHD and CCMD	33	26	8.7	1.4
4	Dirt road with grass MF-MDR	Sunny with bright blue sky, yellow cast BBLD and CCHD	25	22	9.1	3.2
5	High way with overpass LF-HDR	Sunny with bright blue sky and shade under the overpass BBMD and CCMD	23	14	4.1	1.5
6	Dirt road in residential area HF-HDR	Partly sunny, strong yellow cast in front image BBLD and CCHD	27	25	**11.4**	**1.5**
7	High-way with yellow lines LF-LDR	Partly sunny, yellow cast in front and rear image BBLD and CCLD	21	17	5.9	1.7
8	Garage entrance MF-HDR	Overcast with snow, inside yellow illumination BBHD and CCMD	26	16	6.2	4.7

ation of colour hue around the estimated LCM value and as such it is used to provide information on the colour hue uniformity and also about the confidence level that this metric is reliable for this particular corner image. If the LCV value is high, it indicates that either the LCM value is not reliable enough or that the mean colour hue has been removed but still the colour harmonization performance is not of sufficient quality level. Therefore, in such a case we reduce the influences of the corner textures, with high LCV, to the final colour correction metric computation. This final colour correction metric (CCM) is determined as a weighted average of LCM values of all corner textures, where weighting is inversely proportional to the corresponding LCV value.

Experimental results

The proposed test system and methodology for surround view performance evaluation is verified using previously recorded multi-camera sequences, as previously described, with different 3D content and environment conditions. We present results for 8 selected sequences for the evaluation of the brightness and colour harmonization (no low light cases), as described in Table 1. The selected 3D content contains city road and high way, parking places with and without surrounding vehicles and grass, dirt road and garage. The same 3D content is also described in terms of the spatial frequency and dynamic range, as well as in terms of conditions, level of brightness and colour harmonization distortion.

In Table 1, we present results for brightness and colour harmonization evaluation, for the selected sequences with the 3D content and environment conditions. In the column *conditions*, we also provide information of severity of harmonization degradation in terms of brightness and colour. This practically denotes subjective opinion/score about the surround view visual quality perception, which should optimally be represented through BB and CC metric. Note that in this experiment we have only performed subjective scores, related to brightness colour harmonization, using 3 classes: 1) low level of distortion (LD), 2) medium level of distortion (MD) and 3) high level of distortion (HD).

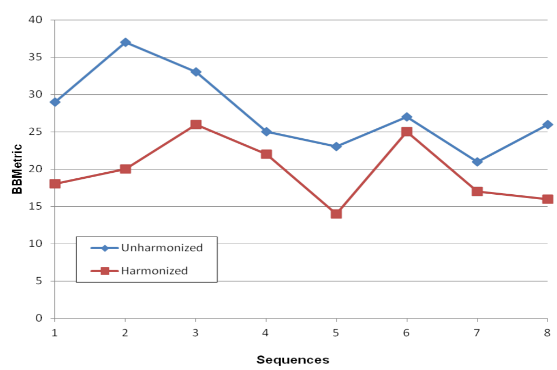

Figure 7. Graph representing brightness harmonization objective metric for the un-harmonized and harmonized surround view images.

The correspondence of the subjective visual quality difference between the harmonized and un-harmonized image textures, depicted in Fig. 3, to the corresponding BB metric values (provided in Table 1), in case of sequence No.2, demonstrate good performance of the proposed BB metric. Further on, in Fig. 7, we present graph depicting the difference between the BB metric for all harmonized and un-harmonized sequences. The results show relatively good performance of the proposed metric, which are shown to be relatively robust to different dynamic ranges of sequences with small dependence on spatial frequency content.

Finally, in Fig. 8, we show colour harmonization evaluation results for the harmonized and un-harmonized sequences, related to data presented in Table 1. As can be seen from Fig. 8, the CC

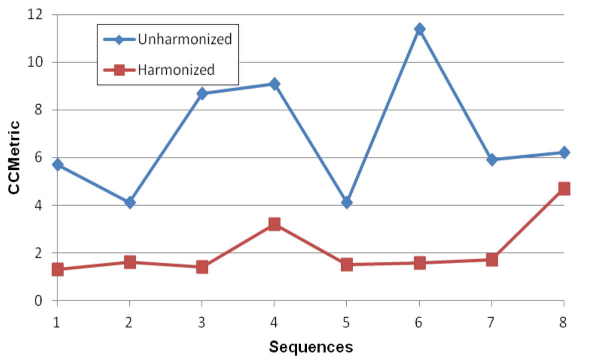

Figure 8. *Graph representing colour correction/harmonization objective metric for the un-harmonized and harmonized surround view images.*

metric produces relatively robust results for different image content and colour cast, except in case of sequence number 4 ("Dirt road with grass") and sequence number 8 ("Garage entrance"). In case of sequence 4 it is related to complex dirt road structure and its surrounding, while in case of sequence 8 it is related to multiple illumination sources (indoor with yellow cast). However, even in those cases the CC metric provides large enough discrepancy values between the harmonized and un-harmonized cases, as can be visually confirmed by example shown in Fig. 5 and its corresponding CC metric values, provided in Table 1 in case of sequence No.6.

Conclusions and Discussions

In this paper we have proposed a novel framework and test system design for more systematic and automatic evaluation of the visual automotive surround view systems. The proposed methodology includes selection of test multi-camera sequences and includes new proposed metrics for objective visual quality assessment of the surround view systems.

References

[1] P. Denny, E. Jones , M. Glavin, C. Hughes, B. Deegan, "Imaging for the Automotive Environment", Handbook of Visual Display Technology, Springer Berlin Heidelberg, Berlin, Heidelberg, pp.1-18, 2015.

[2] Y.-C. Liu, K.-Y. Lin, and Y.-S. Chen, "Birds-Eye View Vision System for Vehicle Surrounding Monitoring", Springer-Verlag, RobVis 2008, LNCS 4931, pp. 207218, 2008.

[3] V. Zlokolica, B. Deegan, P. Denny, M. Griffin, B. Dever, "Free-view multi-camera visualization and harmonization for automotive systems", Electronic Imaging, Autonomous Vehicles and Machines 2017, pp. 12-17(6).

[4] Jaeseok Kim, Hyunchul Shin, "Algorithm & SoC Design for Automotive Vision Systems", Springer Netherlands, 2014.

[5] Dirk W. Hertel, Edward Chang, "Image Quality Standards in Automotive Vision Applications", Proceedings IEEE Intelligent Vehicles Symposium, Istanbul, Turkey, June 13-15, 2007.

[6] F. Dierks, EMVA 1288 - A standard for characterizing cameras.

[7] R. Segur, Using photographic space to improve the evaluation of consumer cameras, in Proc. IS&T PICS, Portland, OR, 2000, pp. 221-224.

[8] W. Kubinger, S. Borbely, H. Hemetsberger, R. Isaacs, "Platform for evaluation of embedded computer vision algorithms for automotive applications", 13th European Signal Processing Conference, 2005.

[9] S. Szabo, R. J. Norcross, J. A. Falco, "Objective test and performance measurement of automotive crash warning systems", Proceedings Volume 6561, Unmanned Systems Technology IX; 65610J (2007), Orlando, Florida, United States.

[10] ITU-T P.910: Subjective video quality assessment methods for multimedia applications.

Author Biography

Vladimir Zlokolica received PhD degree in Applied Sciences at Ghent University, Belgium in 2006, where he also worked as post-doc researcher until 2007. Since 2012 he has worked in Valeo Vision Systems as a Vision Research Engineer and since 2008 he also holds Assistant Professor position at University of Novi Sad, Serbia. His main expertise and interests are in domain of video processing, camera imaging, computer vision and advanced visualization systems.

Mark Griffin received his Bachelor Degree in Applied Physics from University of Limerick in 2010. From 2010 to 2011 he worked as an engineer in ASML, Netherlands. Since 2011 he is working for Valeo Vision Systems, Ireland, as an Algorithm embedded developer in Image Quality Group. His work mostly concerns video algorithm design and development, camera imaging and visualization system development for automotive systems.

Aidan Casey received his Bachelor Degree in Applied Physics from University of Limerick in 2010. From 2010 to 2011 he worked as a test engineer in ASML, Netherlands. Since 2011 he working for Valeo Vision Systems as an Image Quality Function Owner. His work mostly concerns camera imaging and ISP optimization.

Daniela Solera received her Bachelor Degree in Imaging and Photographic Technology from the Rochester Institute of Technology in 2015. Since October 2016 she has been working in Valeo Vision Systems as an Image Quality Engineer. Her work mostly concerns procedure optimization and camera testing.

Brian Deegan received a PhD in Biomedical Engineering from the National University of Ireland, Galway in 2011. Since 2011 he has worked in Valeo Vision Systems as a Vision Research Engineer focusing on Image Quality. His main research focus is on high dynamic range imaging, topview harmonization algorithms, LED flicker, and the relationship between image quality and machine vision. Since 2014 I have been an official Valeo Expert on Image Quality.

Patrick Denny received his PhD in Physics in 2000 from the National University of Ireland, Galway, where he is also an Adjunct Professor of Automotive Electronics. He is a Senior Research Engineer and a Valeo Senior Expert and has worked for the last 15 years at Valeo Vision Systems and its previous incarnation, Connaught Electronics Limited, initially as the Team Leader of RF Design, before moving into the development and innovation associated with automotive vision systems. His research interests include several aspects of automotive vision system image quality, components, algorithmic design, systems and data analytics.

Barry Dever received his Bachelor Degree in Electronic Engineering from Limerick Institute of Technology in 2003. From 2004 to 2011 he worked for General Electric in the development of their Security Applications. Joining Valeo in 2011 he has lead teams in the research and development of high precision Active Alignment platforms and real time Image Quality tools and is officially recognised as a Valeo Expert in this area since 2014. He currently leads the Valeo Vision Image Quality team from the Research and Development headquarters in Galway, Ireland.