Proceedings of the
2007 IEEE Intelligent Vehicles Symposium
Istanbul, Turkey, June 13-15, 2007

Image Quality Standards in Automotive Vision Applications

Dirk W. Hertel, Edward Chang, *Sensata Technologies, Inc.*

***Abstract*—Digital cameras are increasingly used in automotive applications. As these cameras are integrated into active safety systems, image quality becomes ever more important. The captured image information is limited not only by the sensitivity and signal-to-noise ratio of the image sensor, but also by such image capture conditions as ambient lighting, camera-object distance, and relative camera-object velocity. For example, the detection of a pedestrian at night might be further impeded by the headlights of oncoming traffic. Image quality performance characterizes the ability of the imaging system to capture vital image information under application-typical capture conditions. Characterizing objective image quality performance requires clearly defined image quality attributes and metrics.**

This paper introduces the basic concepts of objective and subjective image quality, reviews existing image quality metrics and standards that have been developed for digital still and video applications, and explores their applicability for automotive uses where image information is interpreted by a human observer or machine vision application. The automotive photospace is introduced as a useful tool to characterize the automotive image capture conditions, and distinguish them from other still and video applications. As automotive imaging becomes more widespread, early standardization of image quality is important. This will enable automotive camera suppliers and the automotive industry to communicate in a common language when specifying imaging systems so that sufficient image quality under application conditions is ensured.

I. Introduction

Automotive video cameras with low-cost CMOS image sensors are increasingly used in automotive vision systems. The applications range from driver assistance with forward and rear view images displayed on a dashboard screen to active safety systems where the video footage is analyzed by software. Driver assistance video aids the driver with additional image information in situations where visibility is poor (night vision) or obstructed (side or rear vision). Even though traffic volume at night is just one-fifth of daytime volume, about 46% of fatalities occur at night [1]. If driver assistance systems can successfully reduce the overall number of accidents occurring, then the total number of fatalities may also be reduced. However, the successful operation of any vision-based safety system depends first and foremost on the camera's imaging performance, its ability to deliver images that are "useful" [2] for the display- or software-based detection of vital scene information. Failure to deliver "useful" images will compromise the function of the safety system and could even lead to accidents with possible serious consequences to the health of the driver and others, with the potential for costly litigation.

The environmental, stress, and reliability test requirements of the automotive industry do not directly address image quality performance. It is not possible to predict image quality performance from basic design specifications such as pixel count, pixel size, bit depth or frame rate. It is also difficult to predict it from characterization parameters such as responsivity, dynamic range, quantum efficiency, spectral range etc. Comparisons based on such parameters are only meaningful under narrowly defined lab conditions [3].

The characterization of image quality performance requires measurement of the characteristic functions of tone reproduction, modulation transfer, and noise. A framework of ISO standard test charts and metrics [4] is available for measuring these functions on digital still cameras, and can be applied to single video frames. This quantifies the peak quality under narrowly defined laboratory conditions and thus predicts the system's image quality *capability*. However, automotive video footage is captured under a wide range of conditions, with e.g. scene illuminance or camera-object distances varying from one extreme to the other [5]. These factors, also referred to as "photospace", lie beyond the control of the system designer but they substantially influence the *performance* of the imaging system [6] and thus determine the *yield* of images that are usable for the intended purpose of detecting image information. For example, the visibility or detectability of a pedestrian at night is impeded by low illuminance levels in the presence of strong headlights. The identification of traffic signs is impacted by distance and motion blur. A camera-based active safety system can only perform reliably if the imaging yield under photospace conditions meets or exceeds some minimum quality criterion.

Initial comparisons of single frames and video taken by automotive cameras from different manufacturers indicate that despite being qualified for automotive use, their image quality performance differs significantly, especially under critical photospace conditions. The lack of comparable image quality standards for automotive imaging makes comparisons impossible and product decisions difficult. This underlines the need for a standard-based image quality testing and performance evaluation system that enables automotive camera manufacturers, image analysis software developers, and the automotive industry to develop product

Manuscript received January 15, 2007.

D. W. Hertel is with the Vision Sensing Dept. of Sensata Technologies, Inc., Cambridge, MA 02138 USA (phone: 617-715-6619; fax: 617-661-7601; e-mail: dhertel@sensata.com).

E. Chang is with the Vision Sensing Dept. of Sensata Technologies, Inc., Cambridge, MA 02138 USA (e-mail: echang@sensata.com).

specifications and thus make qualified comparisons. For this to happen, major players in automotive imaging must work together to leverage existing image standards, test methods and image evaluation techniques in order to develop an image quality testing program and metrics, and to identify requirements for new image quality standards specific to automotive cameras. This could be modeled on an initiative that was recently launched by the International Imaging Industry Association (I3A) to improve the image quality of camera phones [7].

This paper develops the methodology to measure image quality performance for automotive vision applications.

II. The Automotive Photospace

Photospace is a quantitative description of the image capture conditions. The term was first introduced by Eastman Kodak to quantify the picture taking preferences for a particular imaging system. The photospace distribution *PSD(L, d)* is the statistical description of the frequency of image capture as a function of the primary photospace coordinates object luminance level *L* and camera-object distance *d*. Photospace information can be used as a design tool to maximize the image quality for the intended usage conditions [6].

The ability of an automotive camera to correctly render the luminance ratios of natural scenes in both daylight and at night will depend on the image sensor's sensitivity and dynamic range. Since the luminance ratios within natural scenes can be as high as 10^7, this requires image sensor technology that extends the dynamic range, commonly referred to as high dynamic range (HDR).

The ability of a camera to render the scene sharply depends on the lens' focus and aperture, a sharp image only being possible within a limited range of camera-object distances determined by the depth of field at the fixed focus point and aperture. The combination of exposure time and velocity will determine the amount of added motion blur.

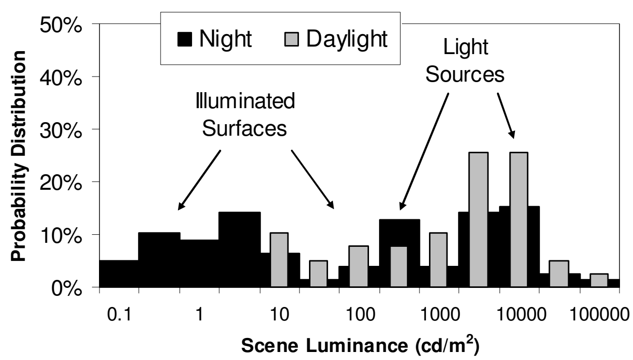

Fig. 1. Initial automotive photospace data with clusters for light sources and illuminated surfaces.

Fig. 1 shows photospace distributions *PSD(L)* measured from a selection of automotive daytime and night scenes. The distributions show clusters that represent light sources (reflections of sun, vehicle lights, street lights, traffic lights) at high luminance levels, and illuminated surfaces (road surface, traffic signs and traffic participants) at the low luminance end of the distributions. The distributions show that the luminance *ratio* within night scenes (10^6) is typically much higher than that of daylight scenes (10^4-10^5), which is due to the much lower luminance of illuminated objects at night.

Figs. 2 and 3 show the imaging performance of different cameras under typical photospace conditions.

Fig. 2. Daylight scene captured with linear (left) and Autobrite [8] HDR image sensor (right). The linear sensor fails to detect the oncoming car in the highlight image area.

Fig. 3. Night scene captured with HDR sensors of different technologies. The Autobrite [8] sensor (right) better distinguishes between different headlights, and between elements of the illuminated road surface.

The obvious differences in imaging performance indicate that camera-based active safety systems would perform differently under critical conditions. This underlines the necessity for characterizing image quality performance.

III. Objective Image Quality

The images from different cameras in typical automotive photospace conditions clearly show the necessity of actually capturing enough image information from the original scene, before it can be viewed and analyzed either by a human observer on a video display, or by feature detection software. Inadequate capture of image information at an initial stage cannot be compensated for at a later stage, and results in decreased photographic yield.

A. The Task of Detail Reproduction

The capture of image information can be described by the model illustrated in Fig. 4 where a large area of the sensor receives a background exposure E, and a small portion of this area receives an additional exposure of ΔE. The two image density levels D_0 and $D_0+\Delta D$, corresponding to the two exposure levels, must be distinguishable for the image information to be retrievable [9]. This requires that the two exposure levels fall within the operating or dynamic range of the sensor where the gradient of its characteristic curve dD/dE is > 0. Signal detection becomes more difficult the smaller the initial signal (exposure difference ΔE) and the lower the gradient. Image noise σ_D leads to a variation of the image densities D_0 and $D_0+\Delta D$ around their mean levels, and

Fig. 4 shows that the broader the overlapping probability distributions of density are, the less the two levels are discriminable. In addition, reproducing detail becomes more difficult as detail size decreases, and will be further compromised by image spread from lens blur and crosstalk.

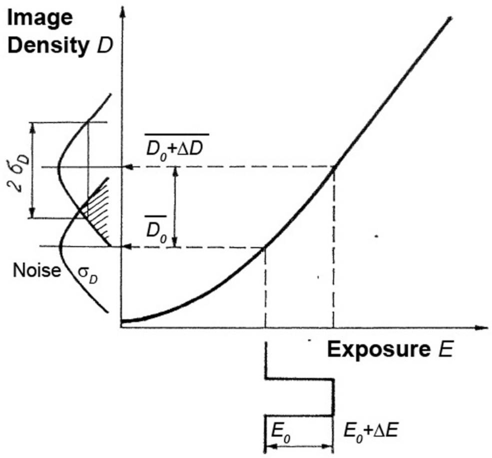

Fig. 4. Detail reproduction: Image signal ΔE transferred by characteristic curve, and its detection compromised by image noise σ_D [10].

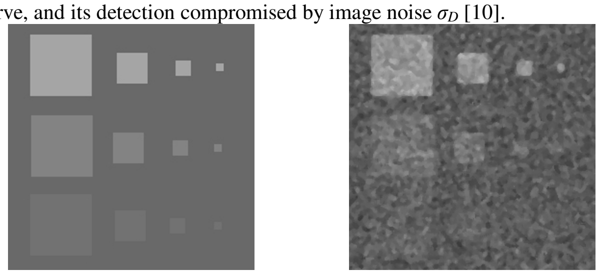

Fig. 5. Small image signals, original (left), and after transfer by a blurry and noisy imaging system (right). The detectability decreases with decreasing signal (top-bottom), decreasing detail size (left-right), and increasing noise.

Fig. 5 illustrates that transfer noise can limit the detection of an image signal that is small in exposure difference and spatial size.

B. *Objective Image Quality Characterization*

The ability of an imaging system to reproduce detail can be fully described using the following three functions [10]: Characteristic Curve of tone reproduction, Noise Power Spectrum (NPS), and Modulation Transfer Function (MTF). These three 'basis characteristics of image quality' are measurable and can also be calculated from structural parameters using statistical models [11]. They describe the ability of the imaging system to reproduce locally variant luminance distributions. Originally defined for photographic image sensors, they have been successfully applied to the characterization of digital imaging systems, and form the basis of ISO standards for characterizing digital camera image quality.

1) *Characteristic Curve of Tone Reproduction*

The Characteristic Curve, conceived and first measured for photographic materials by Hunter and Driffield in 1880 [12], describes the relationship between exposure and the resulting image value (be it in terms of density or any other image units), characterizing how the imaging system transfers large-area image details of varying luminance. ISO 14524 [13] determines this characteristic from standardized test charts for digital still cameras as the Opto-Electronic Conversion Function (OECF). Parameters that characterize image brightness, contrast, and the reproduction of details in the highlights or shadows of the scene can be calculated from the measured OECF. The OECF characterization of HDR cameras for automotive use presents a challenge since conventional standard reflection test charts only cover contrast ranges up to 160:1. Fig. 6 shows the recently developed transmission test chart that contains 20 test patches simultaneously covering a luminance range of $5 \cdot 10^4$. When this is combined with an integrating sphere illuminator where the peak luminance levels can be adjusted over a range of 10^4 [14], an overall luminance range of over 10^8 can be covered in a measurement series.

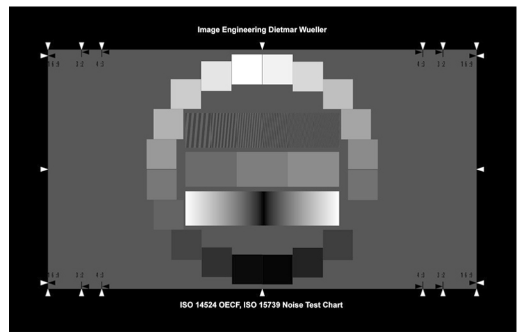

Fig. 6. Transmission test chart for combined measurement of OECF (ISO 14524) and noise (ISO 15739) [14], [15].

2) *Noise Power Spectrum*

Imaging system components such as pixel array or analog image signal processing add unwanted noise to the signal. The NPS measures the local amplitude *variation* of noisy sine waves as a function of spatial frequency. The utility of NPS to describe granularity noise in photographic systems was first demonstrated by Fellgett [16] and Jones [17] in the 1950s. ISO 15739 [15] proposes the simultaneous measurement of NPS as a function of exposure from the OECF test chart. Parameters that characterize sensitivity, signal-to-noise ratio (SNR), dynamic range, and visual noise can be calculated from the measured OECF and NPS curves.

3) *Modulation Transfer Function*

MTF is the frequency-dependence of the change in amplitude of sine waves after transmission by one of the imaging system components such as lens, pixel array, or image signal processing algorithm. Linear modulation transfer allows the cascading of component MTFs in an imaging chain. The utility of the MTF for describing the recording of spatial information on photographic materials was first demonstrated by Frieser in 1935 [9], [12].

ISO 12233 [18] determines the MTF for digital still cameras from standardized test charts. Parameters that characterize image sharpness and resolution can be calculated from the measured MTF.

The MTF characterization for automotive cameras represents a challenge because ISO 12233 uses a high-contrast edge target whereas the most critical features to be detected in automotive scenes are often of low contrast. Recent developments address this issue by proposing a test chart with low-contrast edges [19], see Fig. 7.

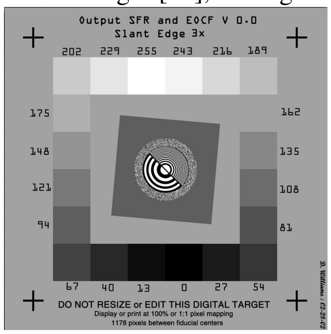

Fig. 7. Reflection test chart with low-contrast edges suggested for the measurement of digital camera MTF [19].

IV. Image Quality Comparison

The ideal comparison criterion would be a single number that characterizes image quality, but due to the complexity of imaging systems such criterion does not yet exist. A direct comparison of the above-mentioned characteristic functions for different imaging systems does not necessarily predict differences in their image quality performance. For example, the better MTF of one camera may be offset by a deterioration of the luminance reproduction (OECF). There are two concepts of calculating comparison criteria from measured characteristic functions [10]: The first one uses the informational SNR as the criterion that characterizes the system's capability to detect information. The second one calculates integral image quality criteria from the basis functions, using specific weighting functions that model the processing of image information by the human observer or detection software.

A. *Criteria of information detection*

The ability of an imaging system to detect small, low-contrast image details is quantified by the photographic SNR q, which is defined in general by the quotient of the image density difference ΔD (signal S), and the standard deviation σ_D of the density fluctuation (noise N):

$$q = \Delta D \big/ \sigma_D \approx \sqrt{S/N} \tag{1}$$

Biedermann and Frieser [20] approximate the photographic SNR for small signal modulations m_0 from the measured basis functions of image quality

$$q(f,D) \approx \frac{\gamma(D)\cdot m_0 \cdot M(f,D)}{\sqrt{N(f,D)}\cdot f}\,, \tag{2}$$

where $\gamma(D)$ is the gradient of the characteristic curve, $M(f,D)$ the MTF, and $N(f,D)$ the NPS at a given image density D. This can be utilized to estimate ranges of exposure (dynamic range) and spatial frequency (detail size) where q exceeds a minimum detection threshold of 1. Higher thresholds are expected to be necessary to ensure the reliability of applications such as feature detection algorithms.

B. *Integral image quality criteria*

Integral image quality criteria require weighting functions that model the detection of image information. This methodology shall be explained for the viewing and evaluation of the image by a human observer, but can also be applied to software-based image interpretation. The methodology requires image attributes and metrics that calculate correlates of these attributes from the measured basis characteristics. The image quality performance is then estimated by photospace weighting of the metrics' values.

1) *Image Quality Attributes*

The goal for automotive image quality can be seen as a balance between naturalness and usefulness, see Fig. 8 [2]. The former strives to produce images with clearly recognizable features, and the latter manipulates images to maximize discriminability of those features presented in the image. For example, a natural image should contain recognizable signal colors or traffic signs. The same image optimized for usefulness might exaggerate local contrast, colors and sharpness to 'see more' features such as a pedestrian on a dark street.

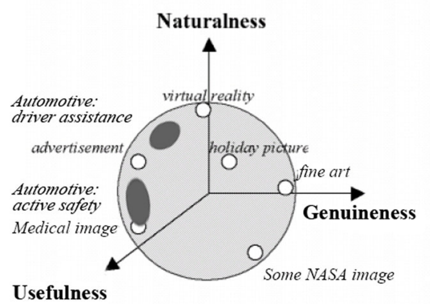

Fig. 8: Image quality goals for different applications (after [2]).

In order to produce a natural image, certain basic expectations of image quality have to be met: The tonal values (luminance ratios) of the original scene must be reproduced so that the overall image brightness matches that appearance of the original scene, and tonal image detail is recognizable. Memory colors, especially those of signals and traffic signs must be reproduced so that they can be recognized unambiguously. Spatial image detail must be reproduced so that it is clearly recognizable without being compromised by image blur or noise.

The image attributes corresponding with these important aspects of image quality expectations are brightness, contrast, color saturation, sharpness and noisiness.

2) *Image Quality Metrics*

Perceptual image quality metrics are designed to correlate with image attributes. They can be calculated from the measured basis characteristics of image quality [21], [22]. These metrics are based on weighting functions that model important aspects of the human visual system such as the contrast sensitivity functions $E(f)$ that describe the visibility of spatial detail in the luminance and chrominance channels as a function of spatial frequency [15], or noise sensitivity functions that describe the relative visibility of noise over luminance and chrominance components of the image [23]. ISO 15739 proposes a visual noise metric that performs contrast sensitivity-weighted integration of the measured NPS over spatial frequency,

$$\sigma_{vis.}{}^{2}(D) \approx \frac{\int \left(\frac{E(f)}{f} \right)^{2} N(f, D) df}{\int \left(\frac{E(f)}{f} \right)^{2} df} . \quad (3)$$

Image quality metrics are important since they reduce the complexity of image quality functions to a single number whilst factoring in important aspects of not only the human visual system but also of image display and viewing conditions. Similarly it would be possible to design metrics that evaluate the usefulness of an image for automatic feature detection by defining weighting functions that model important aspects of the feature detection algorithm.

3) Photospace Weighting

In order to use objective measurements to predict image quality performance under application conditions, the image quality metrics must be applied to a series of test images that sample the application photospace. An example of photospace weighting of sharpness measurements on mobile phone cameras over a wide range of scene illuminance levels is shown in a previous paper [24]. Similarly, the measured distributions of visual noise over photospace dimensions *σ(L, d)* can be weighted with the photospace distribution *PSD(L, d)* to yield photospace averages of visual noise.

V. Experimental

The following example compares the detection capabilities of two automotive cameras. Camera *A* is equipped with a VGA Autobrite [8] HDR sensor, while camera *B* has a wide-VGA sensor with enhanced low-light sensitivity and HDR image processing. Each camera has a Cosmicar 6mm f/1.2 lens.

A. Comparison of two automotive cameras

The ISO 15739 OECF-noise transmission test chart was illuminated with the integrating sphere ETC-LE6-100 [14]. The test images in Fig. 9 were taken at a peak luminance of 1,120cd/m², and the chart modulates the luminance over a range of $5.13 \cdot 10^4$ below the peak luminance. This simulates the imaging of natural scenes that include light sources as well as shadow detail (see Fig. 1). The curves of OECF and SNR vs. luminance shown in Fig. 9 were calculated from the image data, using the recommendations of ISO 15739 [15]. The OECF of camera *A* in Fig. 9 approaches the peak luminance with a steady slope, whereas that of camera *B* is much steeper so that saturation is reached far below peak luminance, leaving the top six patches of the chart almost indistinguishable. At low luminances camera *B*'s SNR is decreased by higher noise, and at high luminances it is limited by saturation. The resulting its dynamic range is reduced by 0.9 log units (18dB) relative to camera *A*.

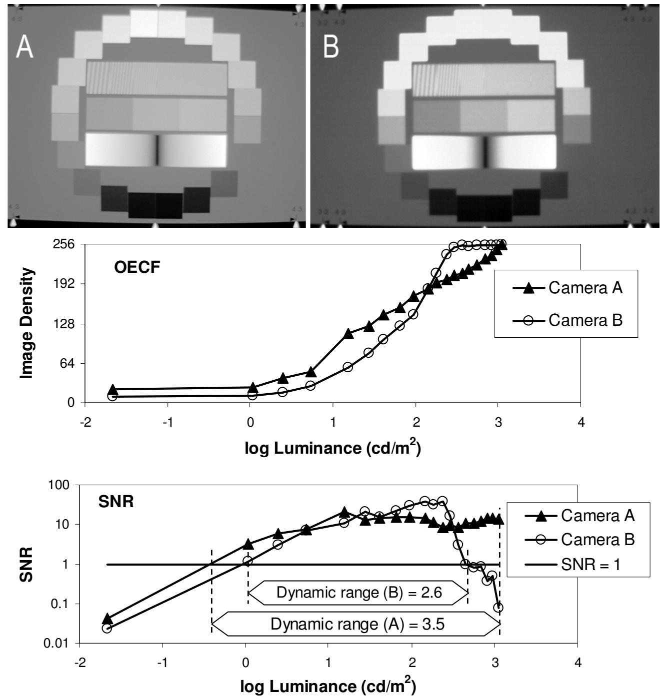

Fig. 9. Images of ISO 15739 OECF-noise test chart and curves of OECF and SNR vs. luminance. Camera *B* (right) shows significant loss of highlight and shadow detail in comparison to camera *A* (left).

In order to simulate a wider range of automotive photospace conditions in the lab, the measurements shown in Fig. 9 were repeated at different peak luminance levels that were varied in logarithmic steps between 4cd/m² and 3,700cd/m². At each of the peak luminance levels, average performance parameters were calculated from the OECF and noise data: visual noise using (3), SNR, and dynamic range. The measurements in Fig. 9 correspond to the data points at 1,120cd/m^2 in Figs. 10 and 11.

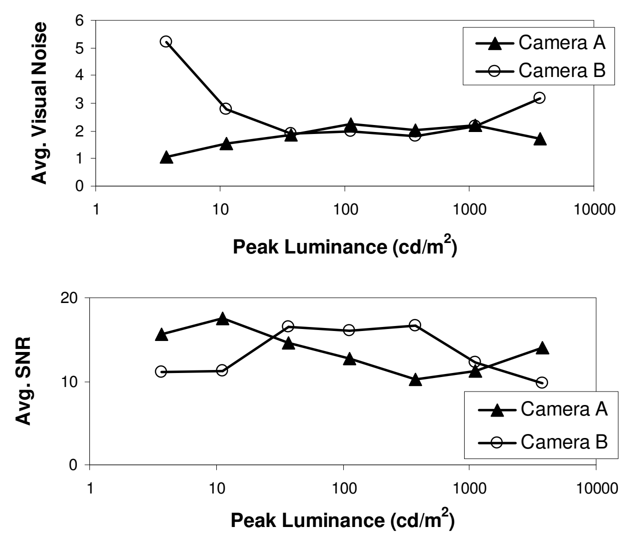

Fig. 10. Average visual noise (top) and SNR (bottom) at different exposure levels.

Fig. 10 shows that both cameras have similar levels of visual noise except for the highest and the two lowest peak luminances where camera *B* shows increased amounts of visual noise. Comparing the average SNRs at the different luminance levels illustrates the interdependence of OECF and noise. At the lowest luminances camera *B* is at a disadvantage due to its higher noise. Its steeper OECF slope gives it an advantage between 10cd/m^2 and 1,000cd/m^2, but at higher luminances the steep slope is responsible for increasing the proportion of saturated highlight levels, with a corresponding decrease in SNR.

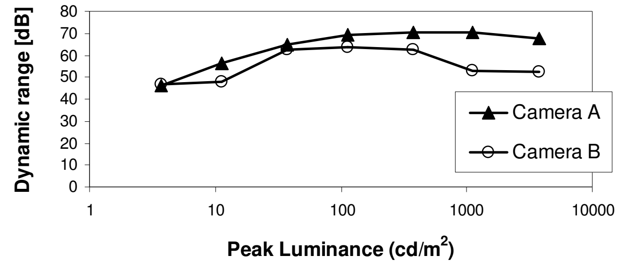

Fig. 11. Dynamic range at the different peak luminance levels.

Fig. 11 compares the dynamic range of each camera at the different peak luminance levels. The higher the luminance, the wider the gap between the dynamic ranges becomes due to the beneficial effect of camera *A*'s Autobrite technology. At low luminances the dynamic range becomes increasingly limited for both by pixel sensitivity. Camera *B* should be at an advantage due to its enhanced low-light sensitivity, but this is offset by higher noise.

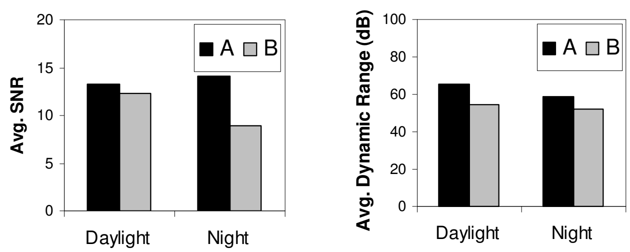

Fig. 12. Photospace averages of SNR (left) and dynamic range (right) for daylight and night conditions.

Finally, photospace averages using the distributions in Fig. 1 are calculated for SNR and dynamic range. Fig. 12 shows that camera *A* has a higher SNR than *B*, and maintains it for both daylight and night conditions over a wider range of scene luminance levels. From the dynamic range data in Fig. 12 it can be estimated that under daylight conditions camera *A* reliably captures image information over a 3.6 times higher luminance range than *B*, and at night its range is still twice as wide.

VI. Conclusion

Image quality performance predictions can be made from measurements of the basis functions of objective image quality, followed by weighted integrations to account for both the image viewing and application-typical photospace conditions. ISO standards for measuring OECF, noise, and MTF form a solid basis for characterizing the image quality performance of automotive cameras, and have been utilized to estimate SNR and dynamic range for automotive photospace conditions at day and night. Future work should concentrate on including object distance and motion in the automotive photospace, and designing application-specific integral image quality criteria that characterize the usefulness of automotive video footage.

References

[1] J. Seekircher, P. M. Knoll, M. Meißner, "The night vision assistant for more driving safety at night," German Future Award 2006. Available: http://www.deutscher-zukunftspreis.de/newsite/2006/team_03/03_material.pdf.

[2] S. Yendrikhovskij, "Image quality: Between science and fiction," in *Proc. IS&T PICS*, Savannah, GA, 1999, pp. 173-178.

[3] F. Dierks, "EMVA 1288 - A standard for characterizing cameras," Available: http://www.baslerweb.com/downloads/9835/EMVA_1288_Praesi.pdf.

[4] ISO (International Organization for Standardization), Technical Committee 42, Working Group 18, Electronic still picture imaging. Available: http://www.i3a.org/wg18.html.

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

[6] B. W. Keelan, *Handbook of Image Quality*. New York: Dekker, 2002, ch. 27.

[7] International Imaging Industry Association (I3A), "What is the I3A Camera Phone Image Quality Initiative (CPIQ)?" Available: http://www.i3a.org/.

[8] H.-S. Lee, C. G. Sodini, K. G. Fife, "Precise MOS imager transfer function control for expanded dynamic range imaging," U.S. Patent 6,600,471, July 29, 2003.

[9] H. Frieser, *Photographische Informationsaufzeichnung*. London New York: Focal Press, 1975, ch. 5.

[10] E. Görgens, "Objective evaluation of image quality," *J. Inf. Rec. Mater.*, vol. 15, no. 5, pp. 305-321, May 1987.

[11] E. Görgens, G. Jehmlich "Electronic-magnetic image recording – image quality estimation for a novel still photography system (in German)," *J. Inf. Rec. Mater.*, vol. 12, no. 5, pp. 283-309, May 1984.

[12] R. Shaw, "A century of image quality," in *Proc. IS&T PICS*, Savannah, GA, 1999, pp. 221-224.

[13] *Photography — Electronic still-picture cameras — Methods for measuring opto-electronic conversion functions (OECFs)*, ISO Standard 14524:1999(E).

[14] D. Wueller, "Evaluating digital cameras," in *Proc. SPIE-IS&T Electronic Imaging, 2006*, SPIE vol. 6069, 60690 K1–15.

[15] *Photography — Electronic still-picture imaging — Noise measurements*, ISO Standard 15739:2003(E).

[16] P. Fellgett, "Concerning photographic grain, signal-to-noise ratio, and information," *J. Opt. Soc. Am.*, vol. 43, no. 4, pp. 271-, 1953.

[17] R. Jones, "New method of describing and measuring the granularity of photographic materials," *J. Opt. Soc. Am.*, vol. 45, no. 10, pp. 799-, 1955.

[18] *Photography — Electronic still-picture cameras — Resolution measurements*, ISO Standard 12233:2000(E).

[19] P. D. Burns, D. Williams, "Refined slanted-edge measurements for practical camera and scanner testing," in *Proc. IS&T PICS*, Portland, OR, 2002, pp. 191-195.

[20] K. Biedermann, H. Frieser, "Über die Informationskapazität photographischer Schichten," *Optik*, vol. 23, no. 1, pp. 75-. 1965.

[21] B. Hultgren, D. Hertel, J. Bullitt, "The influence of statistical variations on image quality," in *Proc. SPIE-IS&T Electronic Imaging, 2006*, SPIE vol. 6059, 60590 O1-10.

[22] D. Hertel, E. Chang, L. Shih, J. Sproul, "Performance evaluation of digital still camera image processing pipelines," in *Proc. SPIE-IS&T Electronic Imaging, 2007*, SPIE vol. 6494, 64940 A1-12.

[23] J. Kuang, X. Jiang, S. Quan, A. Chiu, "Perceptual color noise formulation," in *Proc. IS&T/SPIE Electronic Imaging, 2005*, SPIE vol. 5668, pp. 90-97, 2005.

[24] J. Clark, D. Hertel, B. Hultgren, L. Scarff, "Subjective and objective image quality evaluation," Technical Forum on Camera Phone Image Quality, Stanford University, Palo Alto, CA, Dec. 7, 2006. Available: http://scien.stanford.edu/CPIQ/Presentations.htm