Quality Assessment of Compressed Video for Automatic License Plate Recognition

Anna Ukhanova[1], Jesper Støttrup-Andersen[2], Søren Forchhammer[1] and John Madsen[2]
[1]*DTU Fotonik, Technical University of Denmark, Ørsteds Plads 343, Kongens Lyngby, Denmark*
[2]*Milestone Systems A/S, Banemarksvej 50G, Brøndby, Denmark*

Keywords: Quality Assessment, Video Surveillance, License Plate Recognition, Compression.

Abstract: Definition of video quality requirements for video surveillance poses new questions in the area of quality assessment. This paper presents a quality assessment experiment for an automatic license plate recognition scenario. We explore the influence of the compression by H.264/AVC and H.265/HEVC standards on the recognition performance. We compare logarithmic and logistic functions for quality modeling. Our results show that a logistic function can better describe the dependence of recognition performance on the quality for both compression standards. We observe that automatic license plate recognition in our study has a behavior similar to human recognition, allowing the use of the same mathematical models. We furthermore propose an application of one of the models for video surveillance systems.

1 INTRODUCTION

Video systems and processing are heavily influenced by strong driving forces such as TV distribution, internet applications, mobile communications, and other consumer market products. Surveillance applications are rising in popularity as well. Video surveillance based on IP technology provides advantages compared to classic CCTV systems in terms of cost and flexibility. IP video surveillance is experiencing rapid development and proliferation, but the big markets are oriented on different types of applications. Quality of the video data plays a big role in all of these applications, but in TV-like services video quality relates to visual pleasure e.g. for entertainment rather than for solving specific problems. In video surveillance the ability to recognize objects, persons, events etc. plays a bigger role and thus shifts the perspective of perceived quality.

Quality assessment for recognition tasks is a new assessment scenario which has attracted a lot of attention recently. VQEG (The Video Quality Experts Group) has been driving major research work about consumer video quality in the past years. Recently the group formed a new research project "Quality Assessment for Recognition Tasks" (QART) (Leszczuk and Dumke, 2012) in order to advance task-based video quality research. While there exist well-known and widely used objective quality models such as Peak Signal-to-Noise Ratio (PSNR), Mean Structural Similarity Index (MSSIM) (Wang et al., 2004), Video Quality Metric (VQM) (Pinson and Wolf, 2004) and MOtion-based Video Integrity Evaluation (MOVIE) (Seshadrinathan and Bovik, 2010) for video content in the entertainment sector, there are no well-established models for recognition tasks. While traditional metrics try to define a viewer's overall satisfaction with video quality, quality metrics for video surveillance should define the usefulness of the video data for recognition tasks. Though there is a clear difference between these two types of quality assessment, the new field can definitely benefit from the development of quality metrics for the traditional more entertainment-driven applications.

The resulting quality for a video surveillance system depends on all parts of the signal chain from video capture and compression to transmission, decoding and display in the end-user applications. For example, ambience and environment during capture such as changing light conditions and weather lead to highly varying quality of the video recorded by the cameras. At the same time, transmission errors, such as packet loss, influence final quality as well. While some of these distortions are caused by external conditions such as limited bandwidth or outdoor lighting, artefacts caused by compression can be controlled inside the systems by managing encoding parameters on the camera. Therefore, our work focuses on the

Ukhanova A., Støttrup-Andersen J., Forchhammer S. and Madsen J..
Quality Assessment of Compressed Video for Automatic License Plate Recognition.
DOI: 10.5220/0004671203060313
In *Proceedings of the 9th International Conference on Computer Vision Theory and Applications* (VISAPP-2014), pages 306-313
ISBN: 978-989-758-009-3

quality degradation caused by video compression using H.264/AVC and High Efficiency Video Coding (H.265/HEVC) standards. The novelty of this work is in performance evaluation of these two standards from the point of view of their application to recognition tasks, comparison of the accuracy of modeling the experimental data by logarithmic and logistic functions, and the proposal for usage of the latter model in real-life applications.

The remainder of the paper is organized as follows. Section 2 summarizes related works. Section 3 describes the experiment performed for automatic license plate recognition. Section 4 presents the analysis of the results and a proposed application, and Section 5 concludes the paper giving directions for future work.

2 BACKGROUND AND RELATED WORK

The fast growth of video surveillance technologies and the widespread use of surveillance systems in transportation, law enforcement, etc. have increased the attention to the issues of video quality in such systems. The traditional Quality of Experience (QoE) concept has to be taken differently in surveillance perspectives as task-based applications have different functions from entertainment video. In task-based scenarios it is more appropriate to speak about *Quality of Usefulness* that defines the potential of the video to be used for successful achievement of the recognition task. This is also referred to as visual intelligibility or acuity (Dumke et al., 2011).

Several works have addressed video quality frameworks for recognition tasks in surveillance applications. The Video Quality in Public Safety Working Group was established in 2009 with the support of the Office for Interoperability and Compatibility within the U.S. Department of Homeland Security and the U.S. Department of Commerce's Public Safety Communications Research Program (PSCR). This Working Group has developed a guide for public safety that defines video quality requirements (Video Quality in Public Safety Working Group, 2010). This guide includes definition of some fundamental concepts, introducing a generalized use class concept, recommendations for generalization of use cases into use classes, overview of core video system components, and qualitative guidance for surveillance systems setup. A short summary of the framework proposed in the guide can be found in (Ford and Stange, 2010).

The PSCR project also performed some subjective experiments in order to examine how lighting, target size, and motion together with resolution and bit rate affect the success rate of recognition tasks (Dumke et al., 2011). They did preliminary studies and observed general trends, suggesting further directions in exploring the influence of scene characteristics, bit rates and resolutions on the recognition performance.

Witkowski and Leszczuk (Witkowski and Leszczuk, 2012) applied the framework for describing public safety applications presented in (Video Quality in Public Safety Working Group, 2010) for automatic classification of input sequences into generalized use classes. The proposed method was compared with subjective assessment by humans, and allowed a 70% classification match with end-users opinion. Their analysis led to a conclusion that such automatic classification into use classes has to be additionally verified by humans.

A summary of definitions, research experiments and current trends for quality assessment in surveillance applications is presented in (Leszczuk et al., 2011b). In comparison with other works, this publication describes in addition some standardization activities and discusses general ethical issues.

License plate recognition (LPR) tasks have been addressed in several works as well. Leszczuk et al. describe in detail their subjective experiment on the LPR task (Leszczuk et al., 2011a). The goal of the experiment was to test human recognition capabilities by asking non-expert subjects to detect license plates numbers. This work proposed a simple mathematical model (it was called logit though the formulas represent a logistic model, being inverse to logit) showing the dependency between detection probability and bit rate for a group of test sequences used in the experiment. The fit of this model became less evident when all test sequences were combined together.

Another study (Leszczuk, 2011) presented a case of assessing quality of compressed task-based video on the examples of surveillance videos (LPR scenario) and medical videos (bronchoscopic diagnosis). Test data from (Leszczuk et al., 2011a) was used for analysis for the LPR case and this work suggested modeling of the video quality using a logarithmic function. This study stated that 100% successful recognition could be expected for bit rates higher than 350 kbit/s according to the model, however we would like to note that this number depends highly on the original characteristics and resolution of the video sequences as well as the algorithm used for compression.

Studies (Leszczuk et al., 2011a) and (Leszczuk, 2011) have been further developed in (Leszczuk, 2012). Processed video sequences were grouped into several sets defined by parameters applied to the source video sequences and extended models based on logarithmic and logistic functions were proposed.

Janowski et al. (Janowski et al., 2012) studied both visual (i.e. performed by humans) and automatic LPR (ALPR) using experimental data given in (Leszczuk et al., 2011a). They used two ALPR algorithms for recognition task: Labeling and Artificial Neural Networks and Periodic Walsh Piecewise-Linear Descriptors . These algorithms provided poor performance compared to visual LPR. Analysis of the results showed that in some cases ALPR algorithms provide performance that is different from human subjects, and human subjects easily outperform ALPR algorithms in recognition rate. However, different results may be obtained if other recognition algorithms are applied for ALPR. Automatic extraction of text data in images and video is a challenging problem itself. A review of various approaches addressing this problem is given in (K. Jung, 2004).

In this work we applied two compression algorithms - H.264/AVC (Wiegand et al., 2003) and H.265/HEVC (Sullivan et al., 2012) - on test sequences used in other studies (Leszczuk et al., 2011a). We did our experiment for ALPR and evaluated the performance of the recognition probability models.

3 QUALITY ASSESSMENT EXPERIMENT

In order to develop a quality model, it is necessary to define the ground truth of the video quality through assessment experiments. For task-based applications such as LPR, these experiments can be done either by using evaluation involving human subjects, or automatic (machine) recognition. Below we provide a description of our experiment for ALPR, where we use the recognition probability as a quality measure.

3.1 Source Material

We used 20 video sequences provided by AGH University, Poland (Leszczuk et al., 2011a), and available at CDVL (The Consumer Digital Video Library). They created this set of sequences for their license plate recognition experiments (Janowski et al., 2012; Leszczuk, 2012; Leszczuk, 2011; Leszczuk et al., 2011a). The videos show cars entering and leaving a parking lot. The sequences were recorded directly from a camera with the best possible quality, though some high quality initial compression was performed in the camera. All sequences have resolution 1280×720 pixels and frame rate of 25 fps, with varying number of frames from 479 to 512 resulting in video sequences of approximately 20 seconds length. Further details about sequences recording can be found in (Leszczuk et al., 2011a). Though the car license plate is visible in each sequence for 17 seconds minimum, we have used subsets of frames (from 10 to 310) where the license plate was in a stable position (car stopping before entrance or exit) and in a fixed location in the frame.

We also produced downscaled versions of these sequences with 640×360 resolution using a Lanczos filter. Same subsets of frames were used for the sequences with reduced resolution. The use of downscaled sequences in the ALPR experiment can help to understand how reduction in resolution affects the quality from a recognition perspective.

3.2 Experimental Design

All sequences were encoded with two compression standards - H.264/AVC, using the x264 software implementation (H.264/AVC codec), and H.265/HEVC, using the HM reference implementation (H.265/HEVC codec). The motivation for obtaining results with these two standards was mainly practical, as H.264/AVC is currently used in many video surveillance cameras, while H.265/HEVC promises to be an alternative solution for the cameras in the future.

H.264/AVC was used in Main profile, while H.265/HEVC worked in Low-delay profile. H.265/HEVC is generally more complex and we anticipate that Low-delay profile may be used in initial camera surveillance applications utilizing the new compression standard. More detailed information about the compression parameters for the sequences including the Quantization Parameter (QP) and intra period values is given in Table 1. H.265/HEVC was used with a larger coding unit size as this is one of the main novel features of this standard.

For recognition we used the Intrada ALPR software provided by Q-Free company (Intrada ALPR Software). Its ALPR engine combines a wide range of image processing techniques and statistical analysis with machine learning technology like neural networks. After encoding and decoding video sequences (in YCbCr 4:2:0 format), we extracted the intra frames that fall within the subset of frames defined so that the license plate was in a stable position, as mentioned above. These single intra frames were converted into raw RGB format and given as an input for the ALPR software which operates on uncompressed data. Additionally, some parameters were de-

Table 1: Compression parameters applied for the sequences.

Parameter	H.264/AVC (x264)	H.265/HEVC (HM)
Profile	Main	Low-delay
QP (640×360)	25–51, step of 2	25–51, step of 2
QP (1280×720)	33–51, step of 2	33–51, step of 2
Macroblock / Coding Unit size	16	64
Partition depth	-	4
Intra period	8	8
GOP structure	IPBPBPBP	IBBBBBBB
Slice mode	No	No

fined that include minimum and maximum expected size of the symbols in the license plate and the area within the frame where the recognition should be performed. Car license plates in the test sequences belong to Poland, therefore recognition parameters in the software were set up for this country.

An approximate size of the license plates in the full-resolution sequences is 110×30. We defined one recognition area (366×169 for full resolution, scaled down to 185×85 for reduced resolution) for the video sequences with cars entering the parking lot, and another one (424×170 for full resolution, scaled down to 212×85 for reduced resolution) for the sequences with the cars leaving the parking lot (see Fig.1, area used for recognition is marked by rectangle). As ALPR is typically used in real-time applications, this poses the requirement for quick processing. Operation on intra frames only and predefined region of the frames allows to reduce the processing time.

Figure 1: Example of test sequence (The Consumer Digital Video Library).

For each single input frame the Intrada ALPR software provides the recognized license plate and a so-called confidence level that ranges between 1 and 1000. The confidence level means the trust the ALPR algorithm has in the correctness of the answer given. In other words, the higher the confidence that is reported, the higher the probability that the answer is correct. The reliability of an ALPR algorithm is related to how well the confidence is correlated with the probability of an error. In the Intrada ALPR software the confidence is calculated based on various factors, e.g., how well the detected font and the placement of the characters match the license plate models for the defined country. This allows to discriminate between results that are expected to be correct (high confidence) and results that have less trust.

The used ALPR software is deterministic in its recognition making our results repeatable. In our experiment most of the sequences have more than one intra frame within the defined subset of frames, therefore the ALPR software may provide different answers for these frames. In order to identify the most probable license plate, we evaluated the combined confidence level (*CCL*). *CCL* for each intra frame is computed based on the confidence levels of each character i in the license plate ($CharCL_i$), which are provided by Intrada ALPR, by the following formula:

$$CCL = 1 - \prod_{i=1}^{N}(1 - \frac{CharCL_i}{1000}), \qquad (1)$$

where N is the number of characters in the recognized license plate.

The choice of the final answer for license plate for each compressed sequence in our experiment was made by choosing the license plate with the maximum *CCL*.

4 RESULTS AND DISCUSSION

4.1 Recognition Results and Model

The test set of the experiment consisted of 20 source sequences each encoded with 10 or 14 different QP values depending on resolution (see Table 1), therefore for each compression standard we processed 200 (1280×720) and 280 (640×360) sequences with the ALPR software. The recognition rate for H.264/AVC compressed sequences was 67% (134 out of 200 sequences recognized correctly without a single error) for 1280×720 resolution and 28.2% (79 out of 280 sequences) for 640×360 resolution, while for H.265/HEVC the results were 53.5% (107 out of 200 sequences) and 24.2% (68 out of 280 sequences), respectively. Sequences encoded by H.265/HEVC had on average lower bit rate than H.264/AVC ones.

For each QP value for each compression standard the total amount of symbols was 141 (19 sequences with 7 symbols per license plate and one with 8).

The probability of symbol recognition error was calculated as follows:

$$P_{error} = \frac{N_{error}}{N_{total}}, \quad (2)$$

where N_{error} is a number of symbols recognized incorrectly among all license plates, and $N_{total} = 141$.

Correct symbol recognition probability is defined as:

$$P_{correct} = 1 - P_{error}. \quad (3)$$

Uncoded sequences in full resolution were all recognized correctly ($P_{correct} = 1$), while recognition performance for downscaled sequences was $P_{correct} = 0.68$.

We applied both logarithmic and logistic functions in order to model the recognition probability for ALPR based on either the bit rate or the compression ratio, respectively. The idea of using these functions was initially proposed in (Leszczuk, 2011) and (Leszczuk et al., 2011a) for visual LPR. We extended their experiment by using ALPR software and applying the models for different compression algorithms.

We used the following logarithmic model for modeling the recognition performance:

$$P_{correct} = a \cdot ln(R) + b, \quad (4)$$

where a and b are model parameters that can be obtained by nonlinear regression and R denotes the bit rate. The logistic model is defined as follows:

$$P_{error} = \frac{1}{1+e^{-t}}, \quad (5)$$

$$t = c \cdot C_r + d, \quad (6)$$

where c and d are model parameters that can be obtained by nonlinear regression and C_r denotes the compression ratio, defined as the ratio between the uncompressed size (in YUV 4:2:0 format) and compressed size of the video sequence.

The approximation of the results averaged over 20 sequences for two resolutions for H.264/AVC and for H.265/HEVC are shown in Figs. 2-6.

We used the coefficient of determination R^2 and Pearson correlation coefficient (PCC) to evaluate how well the models fit the experimental data. The achieved R^2 and PCC values are shown in Table 2.

According to R^2 and PCC, the logarithmic model (see a fit in Figs. 2, 3 and 4) makes a good approximation of the experimental data. The logistic model (see a fit in Figs. 5 and 6) in comparison to logarithmic provides higher R^2 and PCC values. Saturation observed on the left and right sides of the plots also supports the preference towards the use of the logistic model.

Table 2: R^2 and PCC for logarithmic and logistic models.

Compr. method (resolution)	model	R^2	PCC
H.264/AVC	logarithmic	0.8372	0.9150
(1280×720)	logistic	0.9910	0.9961
H.264/AVC	logarithmic	0.9578	0.9716
(640×360)	logistic	0.9902	0.9907
H.265/HEVC	logarithmic	0.9403	0.9697
(1280×720)	logistic	0.9725	0.9865
H.265/HEVC	logarithmic	0.9320	0.9563
(640×360)	logistic	0.9852	0.9850

The results for H.264/AVC for the two resolutions (Fig. 2) show that the behaviour of the model depends not only on the bit rate, but also on some other characteristics of the video sequences. After a point around 300 kbps, the reduced-resolution sequences start providing a lower recognition rate compared to full-resolution as the bitrate increases. Between the two test cases for H.264/AVC, only the resolution has been changed which means that the loss of the information caused by downscaling has determined the parameters a and b in the model. These values could depend on sharpness, entropy of the data, or some other video characteristics like amount of high frequences in the video signal which are typically used to ensure reliable recognition of the text on license plates.

Figure 4 shows that the use of H.265/HEVC allows to increase the probability of correct recognition for the same bit rates. As this work studies the influence of compression on recognition performance, we are interested in a wide range of bit rates. In real surveillance systems substantial compression leading to significant quality degradation would most likely not be used. As expected, H.265/HEVC has a better performance if compared at a given bitrate. However, the comparison of recognition rate dependence on PSNR showed that for a range of PSNR values the H.264/AVC provided better performance at a given PSNR (see Figs. 7 and 8). This may be explained by the fact that allowing a significant reduction of the bit rates, H.265/HEVC provided at the same time a lower quality from a recognition point of view.

The results of the experiment also confirmed that ALPR algorithms can be used along with visual LPR by human subjects. The qualitative behaviour of the ALPR is similar to the visual one reported in (Leszczuk et al., 2011a), and our experiment shows that ALPR can provide high recognition rate as well. It supports the idea that ALPR can substitute visual LPR if good ALPR algorithms are used.

It is necessary to note that there is a difference in achievable ALPR performance between compressed and uncompressed video as compression typically affects the high frequencies in the image. Compression can though work as noise filtering as well, which can be beneficial for automatic recognition. The dotted horizontal lines in Figs. 2 and 3 shows the recognition performance for uncompressed sequences for 640×360 resolution ($P_{correct} = 0.68$). Results demonstrate that for some cases a small amount of compression allows to achieve slightly higher average recognition performance than the one for uncompressed sequences.

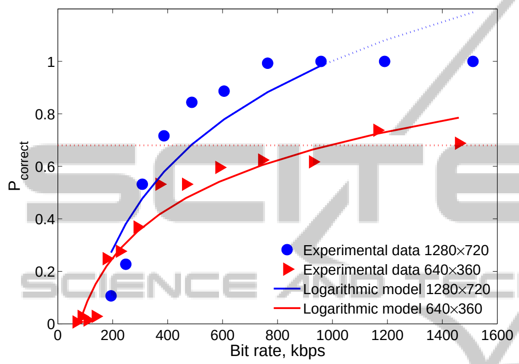

Figure 2: Logarithmic model for H.264/AVC.

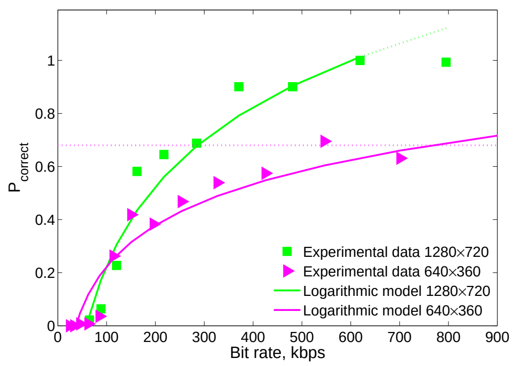

Figure 3: Logarithmic model for H.265/HEVC.

4.2 Proposed Application

We propose a practical application based on the model shown in Figs. 5 and 6. This model shows the dependence of the error probability on the compression ratio. The proposed model can be used for system calibration. Once the surveillance cameras are installed, it is possible to create several test sequences and build the model after performing compression, recognition and analysis of the statistics of the correct recognition. This model is influenced by some characteristics of the scene, such as the size of the characters and

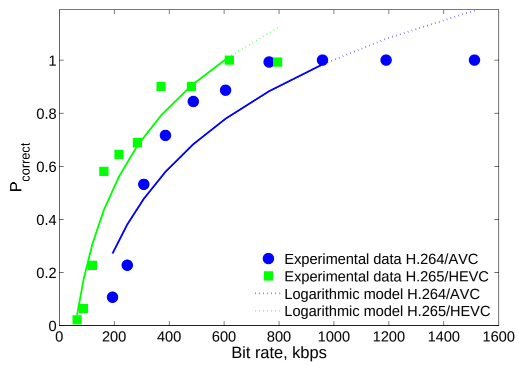

Figure 4: Logarithmic model for H.264/AVC and H.265/HEVC for 1280 × 720.

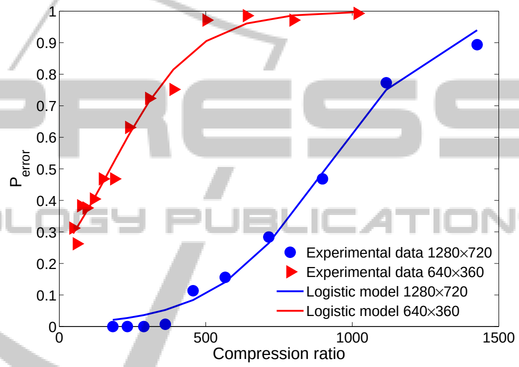

Figure 5: Logistic model for H.264/AVC.

distance of the license plate from the camera, lighting conditions, and camera focus. It is also possible to define a set of different models covering different situations (e.g. in the case with varying lighting for outdoor scenarios).

A logistic model created for a particular surveillance system allows to calibrate the bit rates used for compression with a particular video coding standard in order to provide the acceptable recognition rate. As most practical uses of digital video for surveillance

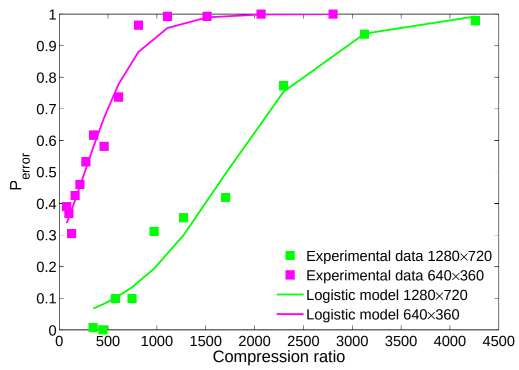

Figure 6: Logistic model for H.265/HEVC.

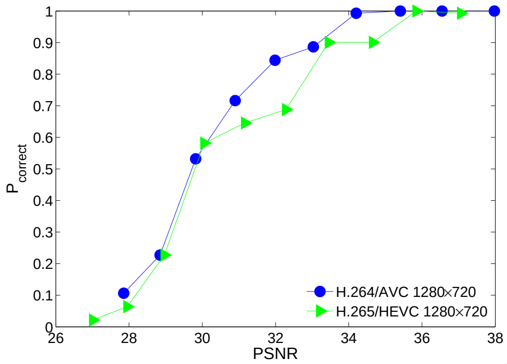

Figure 7: Recognition probability vs. PSNR for 1280×720.

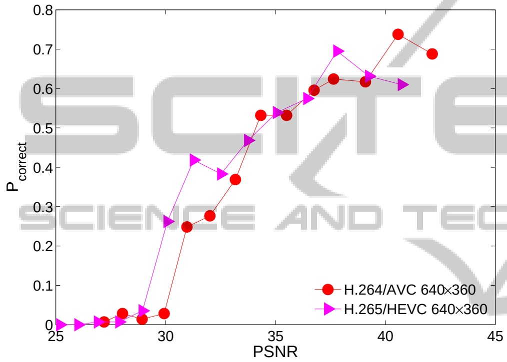

Figure 8: Recognition probability vs. PSNR for 640×360.

require transmission of the video data (e.g. IP-based systems), such a calibration model can also alert the user about the increase of error probability in case there is not enough bandwidth for transmission at the defined bit rate. This means the user could anyway receive the video signal, but would be aware about the decreased reliability of the data. The model can be used for deciding the bit rate for data storage as well. In this case, it is possible to compress necessary frames with a sufficient quality for reliable recognition.

5 CONCLUSIONS AND FUTURE WORK

In this work we have presented the results of an experiment for automatic license plate recognition (ALPR) for evaluating the influence of compression on the quality of the recognition task. We have demonstrated that in this test ALPR has shown similar qualitative behavior to visual LPR and, therefore, can be described using the same mathematical functions.

Our ALPR test using H.264/AVC and H.265/HEVC revealed that their performance can be represented by common recognition probability model but with different parameters. From these results it seems that the model can be used for different compression schemes, though it is premature to conclude that based on the limited data set used for the experiment and since H.264/AVC and H.265/HEVC are rather similar in many aspects.

Coefficient of determination R^2 and Pearson correlation coefficient have shown that the logistic model has higher similarity with the experimental results and therefore it is preferable to use this model.

We performed analysis on the intra frames only and we assume that it is enough to just use the intra frames of the sequence as they tend to have a higher quality compared to P or B frames. This however puts some limitations on the Group of Picture (GOP) size. Too long GOPs lead to long periods between intra frames which could affect the recognition or even lead to the situation where a car license plate appears in the sequence only in the period between two intra frames.

In this work we also proposed a practical application of the logistic model for recognition probability. Such a model can help to calibrate video surveillance systems and choose compression parameters that provide the desired recognition probability.

Our ALPR experiment with different resolutions demonstrated that characteristics of the video sequences play an important role in modeling recognition probability, and in our further works we are planning to extend the described models, identifying the model parameters based on additional information extracted from video data. These experiments together with additional tests in other surveillance scenarios will help defining the Quality of Usefulness for video surveillance applications.

REFERENCES

Dumke, J., Ford, C., and Stange, I. (2011). The effects of scene characteristics, resolution, and compression on the ability to recognize objects in video. *Human Vision and Electronic Imaging XVI, Proc. of SPIE-IS&T Electronic Imaging*.

Ford, C. and Stange, I. (2010). A framework for generalizing public safety video applications to determine quality requirements. *Multimedia Communications, Services and Security*.

H.264/AVC codec. Free software library x264. http://www.videolan.org/developers/x264.html.

H.265/HEVC codec. Reference software HM, version 9.2. http://hevc.hhi.fraunhofer.de/.

Intrada ALPR Software. Q-Free Netherlands B.V. http://www.q-free.nl/.

Janowski, L., Kozlowski, P., Baran, R., Romaniak, P., Glowacz, A., and Rusc, T. (2012). Quality assessment for a visual and automatic license plate recognition. *Multimedia Tools and Applications*.

K. Jung, K. I. Kim, A. K. J. (2004). Text information extraction in images and video: a survey. *Pattern Recognition*, 37(5):977–997.

Leszczuk, M. (2011). Assessing task-based video quality - a journey from subjective psycho-physical experiments to objective quality models. *Multimedia Communications, Services and Security*.

Leszczuk, M. (2012). Optimising task-based video quality - a journey from subjective psychophysical experiments to objective quality optimisation. *Multimedia Tools and Applications*, pages 1–18.

Leszczuk, M. and Dumke, J. (2012). Quality assessment for recognition tasks (QART). *EMERGING 2012, The Fourth International Conference on Emerging Network Intelligence*, pages 69–73.

Leszczuk, M., Janowski, L., Romaniak, P., Glowacz, A., and Mirek, R. (2011a). Quality assessment for a licence plate recognition task based on a video streamed in limited networking conditions. *Multimedia Communications, Services and Security*, pages 10–18.

Leszczuk, M., Stange, I., and Ford, C. (2011b). Determining image quality requirements for recognition tasks in generalized public safety video applications: Definitions, testing, standardization, and current trends. *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–5.

Pinson, M. and Wolf, S. (2004). A new standardized method for objectively measuring video quality. *IEEE Transactions on Broadcasting*, 50(3):312–322.

Seshadrinathan, K. and Bovik, A. (2010). Motion tuned spatio-temporal quality assessment of natural videos. *IEEE Transactions on Image Processing*, 19(2).

Sullivan, G. J., Ohm, J.-R., Han, W.-J., and Wiegand, T. (2012). Overview of the High Efficiency Video Coding (HEVC) standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(12).

The Consumer Digital Video Library. CDVL. http://www.cdvl.org/.

The Video Quality Experts Group. VQEG. http://www.vqeg.org/.

Video Quality in Public Safety Working Group (2010). Defining video quality requirements: A guide for public safety, volume 1.0.

Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E. (2004). Image quality assessment: From error measurement to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612.

Wiegand, T., Sullivan, G. J., Bjontegaard, G., and Luthra, A. (2003). Overview of the H.264/AVC video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(7).

Witkowski, M. and Leszczuk, M. (2012). Classification of video sequences into specified generalized use classes of target size and lighting level. *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–5.