TECHNICAL REPORT (JUNE 2012): RADIAL DISTORTION IN HYBRID VIDEO CODING

Holger Meuel   Marco Munderloh   Jörn Ostermann

Institut für Informationsverarbeitung
Universität Hannover
http://www.tnt.uni-hannover.de/staff/meuel/

ABSTRACT

Modern hybrid video coding consists of basically three main techniques: motion estimation/compensation, transform and quantization and entropy coding. The process of motion estimation is a critical component in terms of efficiency and runtime. If the motion estimation does not work well, the residual to be coded grows and therefore the bitrate will increase. In the presence of lens distortions like radial distortion, the block-based motion estimation process is degraded: the content of compared blocks used in the motion estimation process is differently distorted, depending on the distance of the employed blocks to the distortion center. Actual coders do not consider lens distortions. In this paper we investigate the influence of radial distortion in hybrid video coding. We explore the achievable coding gain and show the improvement of in-loop radial distortion compensation in the motion estimation process. Furthermore we evaluate possible bitrate improvements by considering the radial distortion compensation in the latest reference software of the upcoming video coding standard HEVC.

Index Terms— Hybrid Video Coding, Motion Estimation improvement, Radial Distortion Correction, In-Loop Radial Distortion Compensation, Theoretical Reflection, HEVC

1. INTRODUCTION

The goal of video coding in general is to reduce the storage size of an arbitrary video sequence while retaining subjectively good to perfect quality, depending on given parameters like quantization parameter (QP) which directly correlates with the bitrate. From a simplified point of view basically three main techniques – namely motion estimation/compensation, transform and quantization and entropy coding – are employed in modern hybrid video codecs like *Advanced Video Coding* (AVC) [1, 2, 3, 4, 5, 6] and the upcoming successor *High Efficiency Video Coding* (HEVC) [7] which currently is in the standardization process run of the *Joint Collaborative Team on Video Coding of MPEG & VCEG* (JCT-VC).

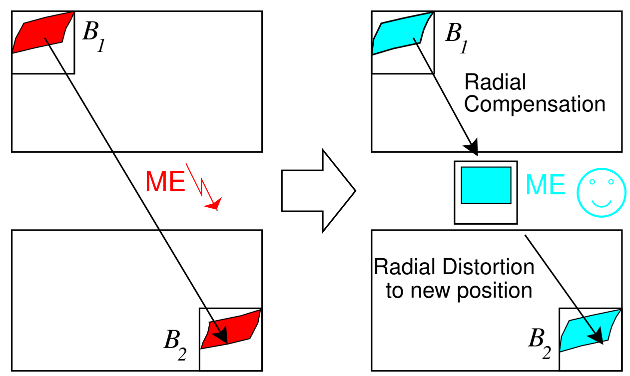


Fig. 1. Principle of radial distortion compensation in motion compensating video coding: the real object in block B_1 moves translational resulting in different image blocks B_1 and B_2. In an in-loop radial distortion compensating system (Figure right) the matching process works better since the shape distortion was adjusted to match the new position.

Motion estimation (and compensation) is one of the basic features of a modern video codec and the efficiency of any hybrid codec depends largely on the quality of the motion estimation process. Although lens distortions like radial distortion do matter even in modern camera systems, it is not considered in video coding at all. Figure 1 outlines the principle of benefiting by considering lens-distortions in motion compensating video codecs: even in the case of a perfect motion estimation, a systematical radial distortion error remains present due to the comparison of two blocks from different positions in the image.

Since the main goal in video coding is to reconstruct the original signal regardless of any “distortions”, any form of preprocessed distortion-correction would inevitably introduce non-reversible errors in the input sequence. Therefore we have to consider the distortion-correction in-loop, which enables us to utilize possible gains from the correction step while the encoder retains the original signal concurrently. After the motion estimation/compensation process the prediction error will – after several other steps like transform- and entropy coding – be transmitted in the bit stream.

In this paper we derive the errors introduced by radial distortion in a theoretical reflection and analyze its impact on the motion prediction and the coding efficiency in a real coding system. The remainder is organized as follows: Section 2 outlines the basics of modern hybrid video coding and points to a principal design problem if test sequences with radial distortion are coded. Section 3 gives a short introduction into the topic of radial distortion in general, presenting the mathematical model of radial distortion used in this paper. In Section 4 we show how we employed radial distortion compensation in a hybrid video coding system. Section 5 describes the improvements achieved for the motion estimation and reflects our implementation in the reference software HM to validate our theoretical findings. Section 6 gives a short summary and an outlook.

2. HYBRID VIDEO CODING

Although there has been a lot of research in the fields of segment- and object-based coding [8, 9] as well as in object-based analysis-synthesis coding [10], the efficiency of the classical hybrid video coding system introduced with H.261 in 1988 [1] is unbeaten. A principle block diagram is printed in Figure 2. It basically consists of three main techniques: first a motion estimation followed by a motion compensation (MC) step is performed. The prediction error is transformed and quantized, often with a DCT based integer transform to decorrelate the spatial signal. Finally, entropy coding is the last important step of modern encoders.

The purpose of the motion estimation/compensation process is to recognize structures similar to structures already known to the decoder and only encode a motion vector pointing to the origin of the pattern and the transformed prediction error, which is called residual. Motion estimation is employed block-wise by comparing the current block to a list of reference blocks and calculating the difference. For complexity reasons often used measures are SAD and SSD, albeit the logarithmic measure PSNR is commonly employed for quality evaluation of coded video sequences compared to their uncompressed original.

3. RADIAL DISTORTION

Radial distortion is one of the main distortions of optical lenses [11]. Referring to [11], [12] found that for many machine vision applications tangential distortion does not need to be considered, thus our reflections only consider radial distortion. [13] states that the correction of the first order of radial distortion is sufficient for an accuracy of about 0.1 pel which is beyond the current subpel interpolation in modern video coding ($\frac{1}{4}$ for AVC and upcoming successor HEVC).

The two different types of radial distortions are shown in Figure 3.

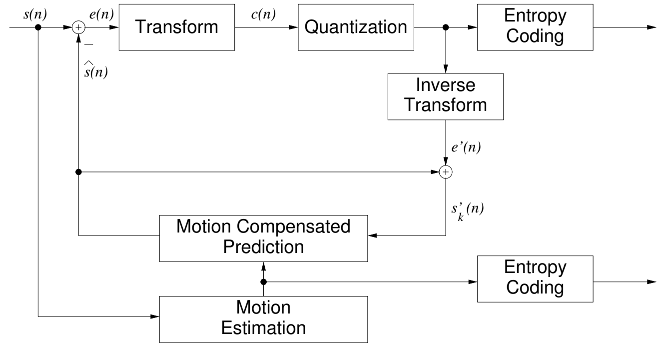
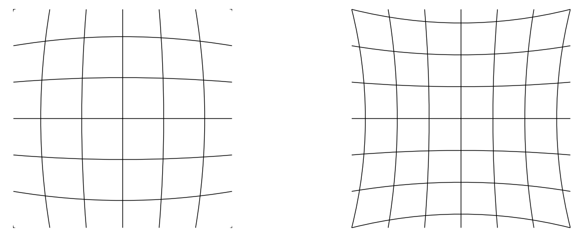


Fig. 2. Simplified block diagram of a hybrid video coder.



(a) Barrel distortion (b) Pincushion distortion

Fig. 3. Possible types of radial distortions.

3.1. Related Work

There has been a lot of research about radial distortion and radial distortion correction [13, 14, 11, 15]. Also in computer-vision radial distortion has to be corrected depending on the special application scenarios [16].

Most correction methods rely on any kind of test pattern to calibrate a lens at a given focal length. In video coding the kind of camera and lens are unknown most of the time and thus all information can only be estimated from the input sequence.

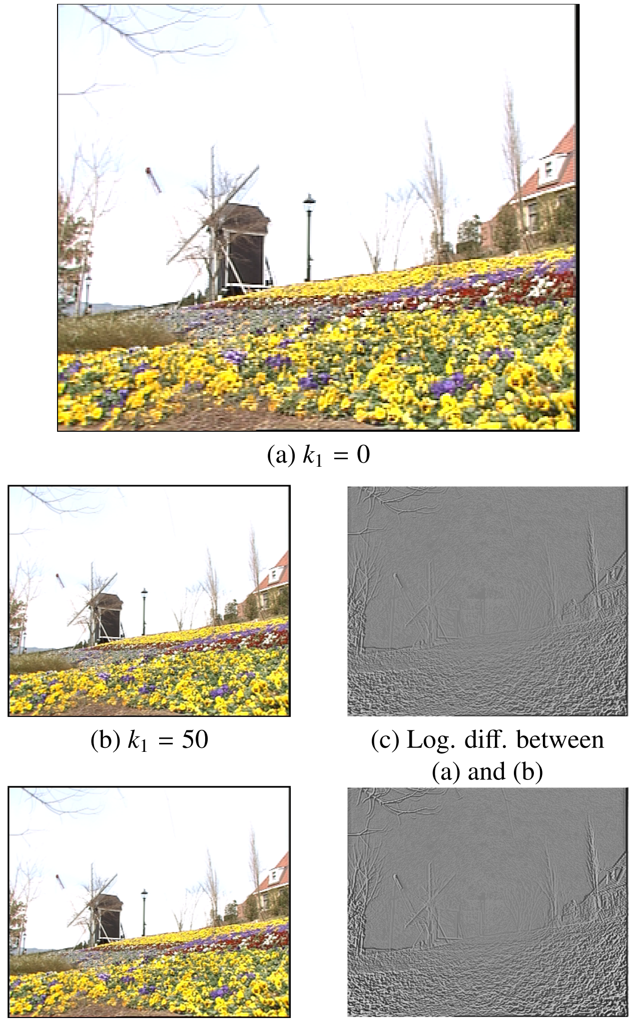
[11] proposed an approach to estimate the radial distortion from the input sequence. Therefor he minimizes the distortion error of different estimated radial distortion parameters iteratively while taking into account the straightness of predetermined lines in the image.

In [17] the complete camera matrix including radial distortion is estimated. The described method is based on the estimation of projective homographies from corresponding image feature points, but works only for static scenes and limited degrees of freedom and thus is not appropriate for video coding.

However, radial distortion compensation has never been employed in video coding to improve the coding efficiency.

3.2. Model of Radial Distortion

The mathematical model of radial distortion is common knowledge and will be summarized shortly [16], [11].



(a) $k_1 = 0$ (b) $k_1 = 50$ (c) Log. diff. between (a) and (b) (d) $k_1 = 100$ (e) Log. diff. between (a) and (d)

Fig. 4. *Flower Garden* (CIF) (artificially) distorted with different radial distortion parameters for illustration of different radial distortion parameters κ_1.

It is assumed that a pinhole camera would lead to non perspectively distorted results. The mapping of a real camera lens can be seen as a homography modeling the deviations to the pinhole camera. Therefore the model projects each image point (x_d, y_d) of the distorted image to undistorted image coordinates (x_u, y_u). For further calculations we use shifted coordinates having its principal point aligned to the center of the radial distortion. We assume it to be in the center of the image; therefore the coordinates are shifted by half of the image width w and height h.

Referring to [11], the image distortion transform can be described by a radial distortion function R. It can be expressed by an infinite series:

$$\begin{aligned} x_u &= x_d\left(1 + \kappa_{1_x} r_d^2 + \kappa_{2_x} r_d^4 + \ldots\right) \\ y_u &= y_d\left(1 + \kappa_{1_y} r_d^2 + \kappa_{2_y} r_d^4 + \ldots\right) \end{aligned} \tag{1}$$

with r_d being the distance from the center of the radial distortion:

$$r_d = \sqrt{\left(x_d^2 + y_d^2\right) \cdot s} \tag{2}$$

The normalization factor s was set to $\frac{1}{w \cdot h \cdot 10^4}$, leading to practical values of $-100 \leq \kappa_{1_x} = \kappa_{1_y} \leq +100$.

As mentioned above, for an accuracy of about 0.1 pel it is sufficient to consider only first order radial distortion in current video coding due to the limited motion vector accuracy [13, 12]. κ_{1_x} and κ_{1_y} are the radial distortion parameters in horizontal and vertical direction, respectively. This general description enables the consideration of oval lenses in principle, although we will assume radial symmetric lenses in our calculations which leads to:

$$\kappa_1 := \kappa_{1_x} = \kappa_{1_y} \tag{3}$$

With this assumption, Equation 1 can be simplified to Equation 4. Given the distorted coordinates $\mathbf{x_d} = (x_d, y_d)$ the undistorted coordinates can be calculated:

$$\begin{aligned} x_u &= x_d\left(1 + r_d^2 \kappa_1\right) \\ y_u &= y_d\left(1 + r_d^2 \kappa_1\right) \end{aligned} \tag{4}$$

Figure 4 displays some examples of different radial distortion parameters for the test sequence *Flower Garden* (CIF) to give an impression of the *strength* of the radial distortion in existing test sequences. In fact, distortions having a radial distortion parameter of $|k_1| \leq 100$ are barely visible.

4. RADIAL DISTORTION IN HYBRID VIDEO CODING

To make use of radial distortion compensation in-loop, we need to remove the radial distortion during the motion search process for the position of the current source block B_1 and reapply the radial distortion for the new position of the destination block B_2 (Figure 1). Figure 5 depicts where the coding gain comes from in principle: even in the case of perfect motion estimation the house in frame $(t-1)$ has a different shape as the house in frame t due to radial lens distortion resulting in a systematical prediction error (filled area in the lower block). After applying radial distortion correction, the shapes of the house looks the same in both blocks, leading to a lower prediction error.

To be able to apply the new radial distortion to the displaced undistorted coordinates, we need to calculate the inverse distortion function. For this we need to compute the distance $\hat{r}_u$ from the undistorted and motion compensated image point (x_u, y_u) to the radial distortion center:

$$\hat{r}_u = \sqrt{\left((x_u + m_x)^2 + \left(y_u + m_y\right)^2\right) \cdot s} \tag{5}$$

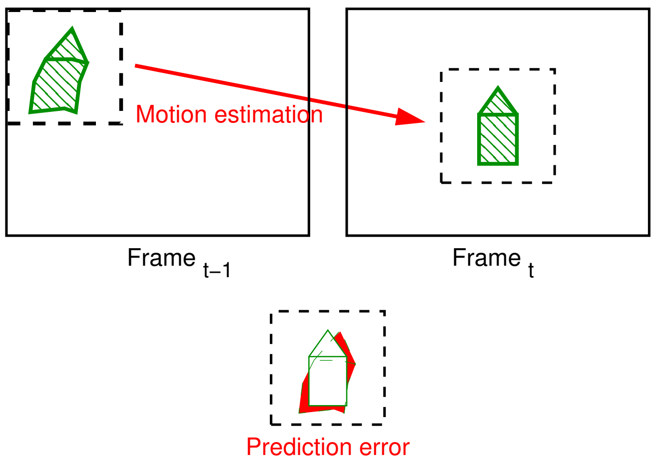


Fig. 5. Principle where the coding gain due to radial distortion compensation in hybrid video coding comes from: the house is different distorted due to radial lens distortion in frames t and $(t-1)$. The resulting systematical prediction error within the motion estimation is depicted in the lower picture.

with $\mathbf{m} = (m_x, m_y)$ being the motion vector.

To get the distorted coordinates $\hat{\mathbf{x}}_\mathbf{d} = (\hat{x}_d, \hat{y}_d)$, the inverse distortion function has to be solved iteratively or in closed form. As in our case the function has to be solved inside the time consuming motion estimation process, we use the following approximation:

$$\hat{x}_d = \frac{x_u + m_x}{1 + (\kappa_1 \cdot \hat{r}_u^2)}$$
$$\hat{y}_d = \frac{y_u + m_y}{1 + (\kappa_1 \cdot \hat{r}_u^2)} \qquad (6)$$

Since the position $(\hat{x}_d, \hat{x}_d)$ is commonly not a fullpel position, an interpolation is necessary to return to the $\frac{1}{4}$ subpel positions of the block matching process. This is done by a bilinear filter on top of the AVC Wiener filter [18].

4.1. Displacement Error Due to Radial Distortion

The influence of the radial distortion increases with the applied displacement size, so the difference between the original distorted block at position $\mathbf{x_d}$ and the radial distortion compensated and displaced block at position $\hat{\mathbf{x}}_\mathbf{d}$ increases for larger motion vectors (which equals to a big frame-to-frame movement). To give an impression of the influence of radial distortion, we assumed a displacement of 16 pel in x and y direction in our following calculation. We examined different radial distortion parameters (1st order) κ_1. Table 1 shows the pixel displacement for the upper left pixel in x and y direction, respectively. Additionally the change of the diameter of the upper left macroblock is given. This diagonal change causes the motion estimation not to work perfectly without any distortion correction. The offsets in column 3 and 4 are the difference between the pixel position of the motion compensated image including a displacement caused by radial distortion and the origin position: $\mathbf{x_d} - (\hat{\mathbf{x}}_\mathbf{d} - \mathbf{m})$. Since one macroblock of 16×16 pel moves nearer towards the image center for small resolutions than for high resolutions (assuming a constant motion vector $\mathbf{m} = (16, 16)$), the displacement caused by radial distortion becomes higher for small resolutions.

Table 1. Pel offset and diagonal change of the upper left macroblock caused by radial distortion for different image resolutions (upper left pixel, motion vector of (16, 16)).

κ_1	Resol.	$\vert$Offset$_x\vert$ [pel]	$\vert$Offset$_y\vert$ [pel]	$\vert$Diagonal Change$\vert$ [pel]
5	1080p[1]	0.0116	0.0067	0.00001
	720p[1]	0.0120	0.0070	0.00002
	SDTV[1]	0.0103	0.0084	0.00005
	CIF[1]	0.0114	0.0095	0.00022
50	1080p	0.0989	0.0569	0.00003
	720p	0.1078	0.0627	0.00005
	SDTV	0.0971	0.0790	0.00031
	CIF	0.1106	0.0927	0.00186
100	1080p	0.1593	0.0918	0.00033
	720p	0.1883	0.1097	0.00031
	SDTV	0.1812	0.1475	0.00013
	CIF	0.2138	0.1792	0.00288

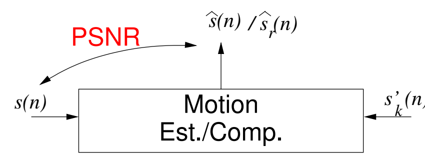


Fig. 6. Comparison between input signal $s(n)$ and motion compensated output once with and second without in-loop radial distortion compensation.

5. EXPERIMENTAL VALIDATION

To validate our theory we measured the accuracy of the motion estimation/compensation. We calculated the PSNR value between the input signal $s(n)$ and the motion compensated prediction $\hat{s}(n)$. The result was then compared with the PSNR measured between the input signal $s(n)$ and the in-loop radial distortion compensated prediction $\hat{s}_r(n)$. We expected to get a better prediction due to radial distortion compensation and hence an improved block matching. Figure 6 illustrates the test set-up. Basically it is an outtake of the hybrid video coder from Figure 2. In any case the output signal is solely the motion compensated image without the prediction error and thus any quantization. If we measure a better PSNR value for our set-up including the radial distortion compensation in

[1] 1080p ≙ 1920 × 1080, 720p ≙ 1280 × 720, SDTV ≙ 720 × 576, CIF ≙ 352 × 288.

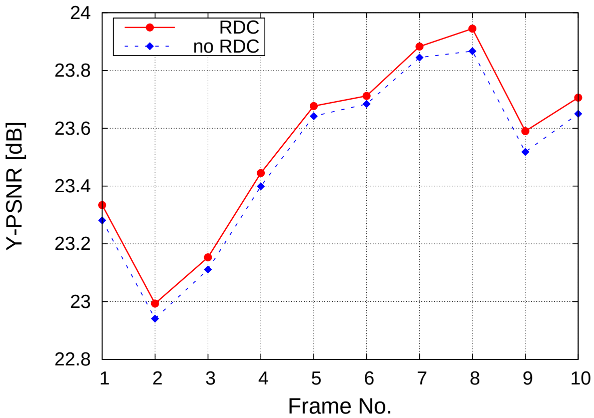


Fig. 7. Frame-wise evaluation of the output of the motion compensated prediction signal.

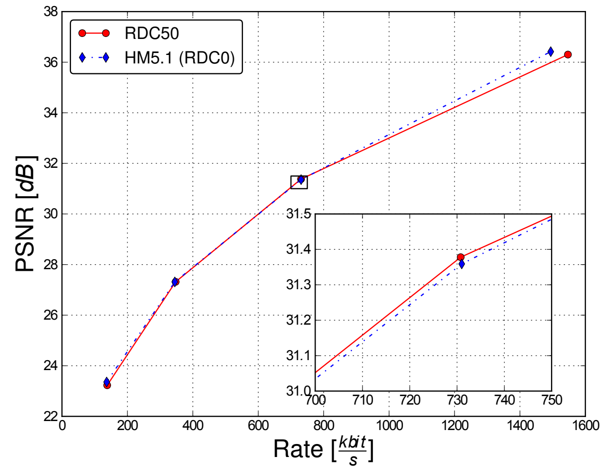


Fig. 8. RD diagram for test sequence *Flower Garden* (CIF) with $k_c = 50$. The solid line corresponds to the proposed algorithm, the dash-dotted line is the HM-5.1 reference.

the motion estimation process we would prove our theory to be true.

Figure 7 shows our findings for several frames of the artificially radial distorted test sequence *Flower Garden*. We observe a better motion prediction for any frame although the improvement is below 0.1 dB. However, the better the motion prediction works, the lower the prediction error becomes resulting in a lower bitrate.

5.1. Implementation in the Reference Encoder

We implemented a radial distortion compensation (correction of radial distortion of 1^{st} order) in the *HEVC Test Model* HM-5.1. We modified the reference block for the motion search in the way described in the previous section.

Since our focus was not on the estimation of the radial distortion, we estimated the correct radial distortion parameter of first order κ_1 manually and determined the optimal *per sequence* value iteratively. We tested the following sequences:

- *Flower Garden* (CIF)
- *Mobile and Calendar* (CIF)
- *Big Ships* (720p)
- *City* (720p)
- Self-recorded sequence with translational movement only (1024×768)

However, our results confirm the calculations from the previous subsection. Figure 8 shows the *Rate Distortion* (RD) *Curve* for *Flower Garden* with QP values between 26 (relatively good image quality) and 44 (fairly bad image quality). We observe small gains at bitrates between 700 and 750 kBit/s in the RD plot of this sequence. Gains of up to 1.2 % based on frame-wise evaluation can be noticed for single frames (P and B), e.g. at a QP of 32 (moderate quality, PSNR about 31.4 dB) and $\kappa_1 = 50$. For the other sequences we tested we got similar results: some frames are up to 10 % better in terms of bitrate when compared with unmodified HM-5.1 coding for CIF sequences. For our 720p sequence *Big Ships* we even got frame rate improvement of up to 12 % for single B frames. Anyhow, this gain can not be caused alone by a better motion estimation since we proved that the pixel displacements caused by radial distortion in the motion estimation process are in or below the range of the motion estimation accuracy. In a highly complex video encoder side effects may lead to better frame coding results initiated by our in-loop radial distortion compensation.

5.2. Discussion of Results

Radial distortion depends on two parameters (Equations 4, 2): the distance from the center of the radial distortion (commonly the image center) and the *strength* of the radial distortion, which depends on physical properties of the lens such as the focal length. We assume to have no prior information about any lens properties. Thus we can only estimate the *strength* of the radial distortion from the video sequence. This *strength* changes with any lens as well as for the same lens if different focal lengths are used. A frame-based radial distortion estimation is out of discussion because in-loop estimation cannot be used for complexity reasons and any transmission of a frame-dependent radial distortion parameter would destroy any gain of the radial distortion compensation.

We have demonstrated that we can improve the motion estimation with our in-loop radial distortion compensation.

However, in a real encoder the pixel displacements caused by radial distortion are in the range of the accuracy limit of modern hybrid video coders. For regions nearer to the center the influence of the radial distortion gets even lower.

6. CONCLUSION

We showed that an in-loop radial distortion compensation can be deployed for improving the motion estimation process.

Based on the well-known theoretical model of radial distortion we derived the deviations for radially distorted images from perfectly non-distorted frames. The maximum pixel offset caused by radial distortion for a feasible radial distortion parameter of $\kappa_1 = 100$ is a about 0.1593–0.2138 pel, depending on the resolution of the input sequence, which lies near the range of the accuracy limit of modern hybrid video coders.

However, our evaluation of a plain in-loop radial distortion compensated motion compensation shows an achievable gain of up to 0.1 dB. We also implemented our approach in the latest development in hybrid video coding, the HEVC reference software HM-5.1. Although coding gains are observed for single frames, the overall bitrate saving is neglectable due to skip modes and the many non-motion compensation related internal encoder optimizations.

7. REFERENCES

[1] ITU-T, *Recommendation ITU-T H.261: Video codec for audiovisual services at p x 64 kbit/s*, Geneva, Switzerland, Nov. 1988. 1, 2

[2] ISO/IEC, *ISO/IEC 11172-2 (MPEG-1 Part 2): Information technology - Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s - Part 2: Video*, Aug. 1993. 1

[3] ISO/IEC and ITU-T, *Recommendation ITU-T H.262 and ISO/IEC 13818-2 (MPEG-2 Part 2): Information technology - Generic coding of moving pictures and associated audio information: Video*, Mar. 1995. 1

[4] ISO/IEC and ITU-T, *Recommendation ITU-T H.264 and ISO/IEC 14496-10 (MPEG-4 Part 10): Advanced Video Coding (AVC)-3rd Ed.*, Geneva, Switzerland, July 2004. 1

[5] Tom Richardson and Rüdiger Urbanke, *modern coding theory*, Cambridge University Press, 1. edition, 2008. 1

[6] Ulrich Reimers, *DVB – Digitale Fernsehtechnik*, Springer-Verlag, 3. edition, 2008. 1

[7] Benjamin Bross, Woo-Jin Han, Jens-Rainer Ohm, Gary J. Sullivan, and Thomas Wiegand, "High efficiency video coding (HEVC) text specification draft 6," in *JCT-VC Document JCTVC-H1003*, 8th Meeting: San José, CA, USA, 2012. 1

[8] Atul Puri, Alexandros Eleftheriadis, and Ros Eleftheriadis B, "MPEG-4: An object-based multimedia coding standard supporting mobile applications," 2003. 2

[9] A. Kaup, "Object-based texture coding of moving video in MPEG-4," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 9, no. 1, pp. 5 –15, Feb 1999. 2

[10] Michael Hötter and Jörn Ostermann, "Analysis-synthesis coding based on the model of planar, rigid, moving objects," *1st International Workshop on 64 kbits/Coding of Moving Video, Hannover*, Jun 1988. 2

[11] Frédéric Devernay and Olivier Faugeras, "Straight lines have to be straight," in *In SPIE, volume 2567*, 2001. 2, 3

[12] R. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *Robotics and Automation, IEEE Journal of*, vol. 3, no. 4, pp. 323 –344, August 1987. 2, 3

[13] H.A. Beyer, "Accurate calibration of CCD-cameras," in *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR '92., 1992 IEEE Computer Society Conference on*, Jun 1992, pp. 96 –101. 2, 3

[14] Thorsten Thormählen and Hellward Broszio, "Automatic line-based estimation of radial lens distortion," *Integrated Computer-Aided Engineering*, vol. 12, no. 2, pp. 177–190, Apr 2005. 2

[15] Tobias Elbrandt and Jörn Ostermann, *Enabling accurate measurement of camera distortions using dynamic continuous-tone patterns*, vol. 18, IOS Press, 2011. 2

[16] R. Szeliski, *Computer Vision: Algorithms and Applications*, Texts in Computer Science. Springer, 2010. 2

[17] Thorsten Thormählen, Hellward Broszio, and Ingolf Wassermann, "Robust line-based calibration of lens distortion from a single view," *Proceedings of Mirage 2003 (Computer Vision/Computer Graphics Collaboration for Model-based Imaging, Rendering, Image Analysis and Graphical Special Effects), INRIA Rocquencourt, France, March, 10-12 2003, pp. 105-112*, 2003. 2

[18] Iain E. G. Richardson, *H.264 and MPEG-4 Video Compression*, chapter 6.5.1.4, John Wiley & Sons Ltd., West Sussex, England, 2003, ISBN: 0-470-84837-5. 4