

# Decoding Complexity-Aware, Rate, Distortion Optimized HEVC Video Encoding

Thanuja Mallikarachchi, Dumidu S. Talagala, Hemantha Kodikara Arachchi, and Anil Fernando

Centre for Vision Speech and Signal Processing, University of Surrey, United Kingdom.

Email: {d.mallikarachchi,d.talagala,h.kodikaraarachchi,w.fernando}@surrey.ac.uk

**Abstract**—Video content adaptation has become a popular application layer approach to reduce the decoding complexity and the associated energy consumption of video playback. In this context, this paper proposes an encoding algorithm to generate less complex HEVC bit streams with minimal impact to the coding efficiency. The experimental results reveal a -17.27 % average decoding complexity reduction and up to -24.91 % energy reduction for openHEVC based software decoding with only -0.73 dB BD-PSNR loss.

## I. INTRODUCTION

VIDEO decoding and presentation are considered as two of the most energy consuming operations for resource constrained mobile hand-held Consumer Electronic (CE) devices (e.g., smart phones, tablets etc.). Moreover, the popularity of High Definition (HD) and Ultra High Definition (UHD) video contents and the proliferation of mobile video consumption have made video playback an evermore challenge for the mobile CE devices. The energy consumption of a CE device during media playback is tightly coupled with the complexity of the content as well as the codec. Therefore, the recently introduced High Efficiency Video Coding (HEVC) standard [1] sufficiently caters the bandwidth requirements for high resolution videos, yet, at the same time demand significant amount of computational and energy resources [2] for the real-time processing of the complex high frame rate HD and UHD video contents. Therefore, reducing the decoding complexity of HEVC bit streams has become a compelling challenge which needs to overcome, especially in the case of resource constrained CE devices [3].

Traditionally, energy reductions in the video decoding devices are achieved by either improving the efficiency of the radio receiver interface, modifying the decoding operations or by modifying the media content to reduce the complexity of the decoding process [3]. The latter being in the domain of video coding algorithms, consists of simplistic approaches that alter the basic coding parameters such as the Quantization Parameter (QP), frame resolution, frame rate, etc. [4], but have a significant impact on the perceived video quality. More state-of-the-art solutions manipulate the motion compensation filters and the in-loop filtering operations introduced in the HEVC to reduce the decoding complexity [5], [6], or adopt Dynamic Voltage and Frequency Scaling (DVFS) [7] techniques to reduce the decoder's power consumption at the hardware

level. However, in general, the state-of-the-art methods in the literature do not exploit the variations of the intricate complexity levels that exist between different decoding operations to determine the optimum coding parameters that minimize the decoding complexity and thereby the decoder's energy consumption in tandem with the coding efficiency. In this context, this paper proposes a novel encoding algorithm that uses its awareness of the decoding complexity of HEVC coding features to reduce the computational resource requirements of the decoder with minimal impact on the coding efficiency.

The remainder of the paper is organized as follows. Section II performs a comprehensive analysis on the decoding complexity, and distortion parameters which is thereafter utilized in the novel encoding algorithm proposed in Section III. Finally, Section IV and V present the experimental results and the concluding remarks, respectively.

## II. DECODING COMPLEXITY-DISTORTION ANALYSIS

The selection of a coding mode that is both optimum in terms of the decoding complexity, rate and distortion requires the encoder to be aware of the impact of the coding mode on all three parameters for the video content under consideration. Thus, estimating the decoding complexity incurred at the decoder within the encoding loop is crucially challenging. In this context, the generic decoding complexity estimation models introduced in our previous works [8], [9] can be utilized and are integrated within the encoder.

Next, the complex relationship that exists between the decoding complexity, rate and distortion needs to be analyzed and modelled. In this context, the investigations carried out in [10] [11] define the relationship between the bit rate and the distortion, which does not sufficiently characterize the impact of decoding complexity on the coding efficiency. Hence, this section first introduces a novel approach followed in this work to analyze the relationship between video distortion and the decoding complexity.

### A. Decoding Complexity-Distortion Model

In order to analyze the relationship between the distortion  $D$  and the decoding complexity  $C$ , the complete parameter space of both parameters must be first observed. An example of the typical decoding complexity-distortion behaviour observed is illustrated in Fig. 1(a). In general, it is observed that the distortion tends to monotonically increase with decreasing

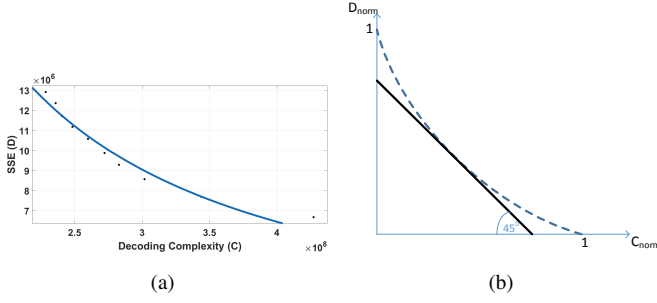


Fig. 1. (a) The typical decoding complexity-distortion behaviour observed for a particular QP. (b) A representation of the normalized decoder complexity and distortion for a particular QP. The depicted tangent to the curve in (b) indicates the  $\lambda_c$  which results in equal trade-off between the decoder complexity and distortion.

decoding complexity, thus, the relationship between the two parameters is therefore best characterized by,

$$D = \alpha C^{-\beta}, \quad (1)$$

where  $\alpha$  and  $\beta$  are content dependent parameters. Thus, for a particular content and coding scenario, the trade-off between the  $D$  and the  $C$  can be expressed in terms of the slope of the distortion-decoding complexity curve, given by

$$\lambda_c \triangleq -\frac{\partial D}{\partial C} = \alpha \beta C^{-\beta}. \quad (2)$$

Expressing the decoding complexity-distortion trade-off as a Lagrangian optimization problem, an optimization function can therefore be formulated as

$$\min_{p \in \mathcal{P}} J_{CD} \quad \left| \quad J_{CD} \triangleq D(p) + \lambda_c C(p), \quad (3)$$

where  $\lambda_c$  represents the Lagrangian parameter. For the general case however, where the relationship in (1) is unknown, a generic  $\lambda_c$  is computed as described next.

### B. Computing $\lambda_c$

Observing the decoding complexity-distortion behavior in Fig. 1(a), it is clear that for a particular QP,  $\lambda_c \in [0, \infty)$  represents a space where decoding complexity and distortion can be traded-off against each other. Thus, the decoding complexity and distortion values that result from an experimental sweep for a range of  $\lambda_c$  values in the optimization function in (3) for  $QP \in \{0, \dots, 51\}$  with different test sequences are further analyzed. Next, the optimum  $\lambda_c$  for a particular QP is computed such that it corresponds to the value of  $\lambda_c$  that results in an equal trade-off of the normalized distortion and the normalized decoding complexity as illustrated in Fig. 1(b). Fig. 2 illustrates the behavior of the optimum  $\lambda_c$  for different values of QP, and the respective relationship is given by

$$\lambda_c = \begin{cases} 8.739 \times 10^{-5} \cdot e^{0.1327 \cdot QP} & \text{Intra frames} \\ 0.001393 \cdot e^{0.09023 \cdot QP} & \text{Inter frames.} \end{cases} \quad (4)$$

The decoding complexity-distortion trade-off modelled above can now be used together with the bit rate trade-off factor to develop a decoding complexity-rate-distortion optimized coding mode selection algorithm as described in the following section.

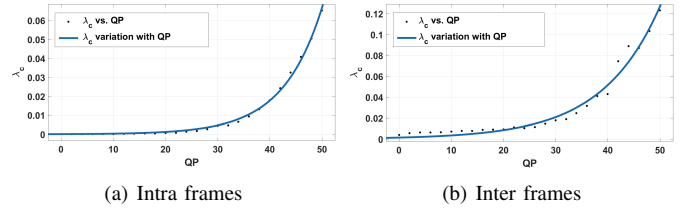


Fig. 2. The  $\lambda_e$  vs QP relationship.

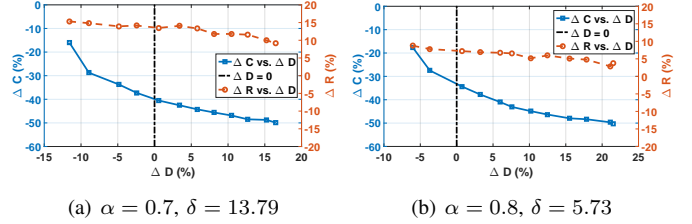


Fig. 3. The typical behavior of  $\Delta D$ ,  $\Delta C$ , and  $\Delta R$  parameters for two different combinations of  $\alpha$  and  $\delta$  values for “Kimono” sequence at QP=30. A curve indicates the variation of respective parameters for a particular  $\alpha$  (i.e.,  $\alpha = \{0.7, 0.8\}$ ) over a range of associated  $\delta$  values. The  $\alpha$  and  $\delta$  values indicated below each sub-figure correspond to the  $\Delta C$  and  $\Delta R$  values at  $\Delta D = 0$ .

## III. DECODING COMPLEXITY-RATE-DISTORTION OPTIMIZED ENCODING

Preparing an encoded bit stream that is optimum in terms of decoding complexity and coding efficiency require an in depth analysis of the impact of the individual parameters. Thus, the Lagrangian cost function for this case can be expressed as,

$$\min_{p \in \mathcal{P}_{opt}} J_{CRD} \quad \left| \quad J_{CRD} \triangleq D(p) + \bar{\lambda}_r R(p) + \bar{\lambda}_c C(p), \quad (5)$$

where  $\bar{\lambda}_r$  and  $\bar{\lambda}_c$  are the bit rate and decoding complexity trade-off parameters, respectively. However, obtaining optimal values for  $\bar{\lambda}_r$  and  $\bar{\lambda}_c$  directly from (5) is not straightforward, and the experimental approach described next is adopted to this end.

First, the general behaviour of the cost function  $J_{CRD}$  is investigated for the range of the two complexity parameters, where  $\bar{\lambda}_r \triangleq \alpha \lambda_r$  and  $\bar{\lambda}_c \triangleq \delta \lambda_c$  for  $\alpha \in \{0, \dots, \infty\}$  and  $\delta \in \{0, \dots, \infty\}$ . Note that  $\lambda_r$  and  $\lambda_c$  in [10] and (3) respectively are used for this purpose. In this study, 50 frames of each test sequence are encoded using HM16.0 reference encoder and the results are used to infer the general behaviour of  $J_{CRD}$  in the full parameter space of  $\bar{\lambda}_r$  and  $\bar{\lambda}_c$ .

Next, in order to determine a suitable operating point in this parameter space, the rate, distortion and decoding complexity at each operating point is compared with the respective values for each obtained when using the rate-distortion Lagrangian cost function in [10]. To facilitate this, the percentage differences of each parameter, i.e.,  $\Delta R$ ,  $\Delta D$  and  $\Delta C$ , given by

$$\Delta \Gamma = 100 \times \frac{\Gamma_{RD} - \Gamma_{CRD}}{\Gamma_{RD}}, \quad (6)$$

is used. Here,  $\Gamma$  represents the distortion  $D$ , bit rate  $R$ , and decoding complexity  $C$ , while  $\Gamma_{RD}$  and  $\Gamma_{CRD}$  correspond to the scenarios where the cost functions in [10] and (5), respectively, are applied.

TABLE I  
SIMULATION ENVIRONMENT

Parameter	value
Encoding configurations	HM16.0 [12], <i>random access</i>
Quantization Parameters	22,27,32,37
Machine	Intel x86 Core i7-6500U
Operating system	Ubuntu 16.04
Decoders	HM 16.0 [12], openHEVC [13]
Decoding complexity measurement	callgrind/valgrind [14]
Decoder's energy consumption	Linux power measurement tools
Frequency governors	Linux ondemand, DVFS [7]

Fig. 3 illustrates a subset of these results for a single sequence. Here, the differences in the behaviour for the different frame types and  $(\alpha, \delta)$  pairings can be observed. It can also be observed that distortion for example deviates significantly from that of the RD optimized value for extreme  $(\alpha, \delta)$  combinations. Next, since one of the objectives of this work is to minimize the impact on distortion, a subspace of the parameter space of  $\alpha$  and  $\delta$  is derived that satisfies the condition  $\Delta D = 0$ . However, from Fig. 3, for the different  $(\alpha, \delta)$  combinations satisfying  $\Delta D = 0$ , variations in  $\Delta R$  values can still be observed. Thus, the operating values for  $(\alpha, \delta)$  are derived by finally minimizing  $\Delta R$  in the subspace of parameters satisfying  $\Delta D = 0$ .

A generic set of scaling factors for  $(\alpha_0, \delta_0)$  are obtained as the average of the parameters obtained for the individual test sequences, and is given by

$$(\alpha_0, \delta_0) = \begin{cases} 0.9, 0.68 & \text{Intra frames} \\ 0.9, 0.33 & \text{Inter frames} \end{cases}. \quad (7)$$

Thus, the rate and decoding complexity trade-off factors in (5) are now become  $\bar{\lambda}_r = \alpha_0 \lambda_r$  and  $\bar{\lambda}_c = \delta_0 \lambda_c$ , respectively.

#### IV. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed algorithm is implemented in the HM 16.0 reference software, where the complexity models in [9] [8] perform the decoding complexity estimations and the proposed Lagrangian cost function performs the coding mode selection. The simulation environment and the measurement tools utilized are illustrated in the Table I.

The performance of the proposed and state-of-the-art algorithms are evaluated by measuring the complexity reduction achieved by the different bit streams at the decoder. To this end, the percentage decoding complexity reduction, and corresponding energy reduction is measured as,

$$\Delta \xi = 100 \times \frac{\xi_\rho - \xi_{\text{HM}}}{\xi_{\text{HM}}}, \quad (8)$$

where,  $\xi \in C, E$  represent the decoding complexity and energy consumption, respectively. In this context, the  $\xi_{\text{HM}}$  and  $\xi_\rho$  correspond to the decoding complexity or energy consumption of the bit streams generated by the HM reference encoder, and the proposed and state-of-the-art algorithms, respectively. The impact on the coding efficiency is measured in BD-PSNR [15] which illustrates the impact on video quality for a given bit rate.

#### A. Performance Evaluation and Analysis

Modification of the motion compensation filters in the decoder (MC) and the intermittent skipping of the loop filter (LF) in the decoder, proposed by Nogues *et al.* [16], contributes significantly to reduce the decoding complexity. However, this impacts visual quality considerably, mainly due to the distortions introduced by the modified motion compensation filtering operations and the propagation of these errors to future frames. In comparison the proposed algorithm suffers from less BD-PSNR reductions due to it being operated on the encoder-side which determines the type of the motion vector (integer-pel vs. fractional-pel) based on the optimization cost function in (5); thus, requiring no changes to the decoding process itself. In addition, skipping the loop filter on the other hand, as in [16], reduces the decoding complexity with minimal impact on quality, and can also be implemented when decoding the proposed bit stream. For example, the experimental results presented in the Table II illustrate the  $\Delta C$  improvements that can be achieved for the proposed algorithm in this manner albeit for a marginal BD-PSNR loss.

The encoding algorithm proposed by He *et al.* [6] attempts to reduce the complexity of the filtering operations during motion compensation (PUM) and the de-blocking (DBLK) performed by the decoder. In comparison to [16], a much lower coding loss is observed, since the required motion vector and filtering decisions are made by the encoder side. More importantly, the decoder complexity-efficient mode selection is only limited to the PU level motion vectors, thus, the  $\Delta C$  achieved is relatively small when compared to the proposed algorithm and [16]. Finally, the trade-off factors being QP agnostic and the lack of a detailed complexity estimation model in its optimization process results in relatively lower decoder complexity reductions. In contrast, the proposed algorithm demonstrates considerable improvements in  $\Delta C$  with minimal impact on the BD-PSNR due to its more comprehensive assessment and selection of trade-off factors for both the bit rate and the decoding complexity. This is aided by the use of a more detailed and accurate decoder complexity estimation model that is based on the HEVC coding features, which yields more accurate decoding complexity estimates for complexity-rate-distortion optimization.

Finally, the observed reduction in energy consumption when using the openHEVC software decoder with different DVFS schemes and bit streams is reported in Table III. Here, when compared to the HM 16.0 generated bit streams, the proposed algorithm demonstrates on average -7.79% decoding energy reduction when using the Linux kernel's ondemand DVFS governor. Thus, the effectiveness of the proposed method's complexity reduction, in terms of the impact on energy consumption, is quite apparently beneficial. Moreover, the effectiveness of using an application specific DVFS governor [7] for video playback is evident in the experimental results shown in the Table III. The complexity reduction by the proposed algorithm's bit streams allow the DVFS algorithm to select much lower CPU operating frequencies that lead to

TABLE II  
DECODING COMPLEXITY REDUCTION PERFORMANCE OF THE PROPOSED ALGORITHM

Sequence	Proposed (Model only)			Proposed (Model + LF [16])			He <i>et al.</i> [6] (PUM + DBLK)			Nogues <i>et al.</i> [16] (MC+LF)		
	$\Delta C^\dagger$ (%)	$\Delta C^\ddagger$ (%)	BD- PSNR	$\Delta C^\dagger$ (%)	$\Delta C^\ddagger$ (%)	BD- PSNR	$\Delta C^\dagger$ (%)	$\Delta C^\ddagger$ (%)	BD- PSNR	$\Delta C^\dagger$ (%)	$\Delta C^\ddagger$ (%)	BD- PSNR
Band	-39.89	-13.59	-0.80	-51.94	-21.99	-1.08	-6.26	-6.84	-0.81	-22.48	-11.56	-0.29
Beergarden	-29.85	-12.53	-0.41	-40.31	-19.28	-0.66	-7.68	-7.15	-0.42	-21.49	-9.94	-0.40
Cafe	-40.69	-13.09	-0.33	-52.42	-20.77	-0.52	-13.82	-9.87	-0.29	-21.22	-11.13	-0.26
Dancer	-40.84	-20.47	-0.89	-49.49	-27.01	-0.64	-21.00	-13.03	-0.68	-34.35	-9.82	-0.46
GTFLy	-43.97	-20.46	-0.86	-54.93	-28.69	-1.04	-21.51	-10.80	-1.33	-36.89	-11.79	-0.83
Kimono	-42.73	-21.41	-0.35	-52.03	-29.72	-0.62	-21.34	-14.95	-0.59	-33.39	-10.52	-0.37
Musicians	-39.75	-18.95	-1.92	-50.02	-26.77	-2.11	-19.09	-10.25	-1.42	-35.86	-10.39	-0.78
Parkscene	-42.65	-20.69	-0.56	-46.68	-27.37	-0.67	-22.07	-13.72	-0.58	-35.56	-9.52	-0.49
Poznan St.	-37.22	-14.30	-0.48	-48.62	-21.58	-0.48	-11.47	-9.41	-0.15	-20.08	-11.42	-0.25
<b>Average</b>	<b>-39.73</b>	<b>-17.27</b>	<b>-0.73</b>	<b>-49.60</b>	<b>-24.79</b>	<b>-0.86</b>	<b>-16.02</b>	<b>-10.66</b>	<b>-0.69</b>	<b>-29.03</b>	<b>-10.67</b>	<b>-0.45</b>

$\dagger \Delta C$  achieved using the HM 16.0 reference decoder.

$\ddagger \Delta C$  achieved using the openHEVC decoder.

TABLE III  
ENERGY CONSUMPTION PERFORMANCE OF THE PROPOSED ALGORITHM

Sequence	Proposed + ondemand $\Delta E$ (%)	HM 16.0 + DVFS [7] $\Delta E$ (%)	Proposed + DVFS [7] $\Delta E$ (%)
Band	-9.36	-12.55	-24.34
Beergarden	-13.14	-22.14	-24.91
Cafe	-10.66	-17.78	-16.96
Dancer	-13.71	-16.60	-30.32
GT Fly	-5.79	-9.26	-16.21
Kimono	-6.61	-16.17	-23.16
Musicians	-6.17	-10.42	-13.21
Park scene	-3.37	-15.84	-23.23
Poznan St.	-1.35	-6.30	-11.26
<b>Average</b>	<b>-7.79</b>	<b>-14.11</b>	<b>-20.40</b>

greater energy efficiency.

## V. CONCLUSION

The increasing demand for high resolution video content and the complexity of encoded bit streams have made the decoder energy reduction a compelling research challenge. To this end, reducing the complexity of the encoded bit stream is seen as a potential application layer solution to reduce the complexities associated with the decoder operations. In this context, this paper introduces an encoding algorithm for HEVC to prepare bit streams with minimal decoding complexity while keeping coding efficiency loss to a minimum. The experimental results reveal decoding complexity reductions of -39.73%, and -17.27% with only -0.73 dB BD-PSNR for the HM 16.0 reference and openHEVC decoders, respectively. In addition, the energy consumption analysis reveals that the proposed algorithm can reduce the device's energy consumption -7.79% for the openHEVC software decoder such as openHEVC. Moreover, utilizing an application specific DVFS governor improves the energy consumption performance of the proposed algorithm further, achieving -20.40% reductions. Future works will focus on introducing decoder energy-aware adaptive video streaming solutions.

## REFERENCES

- [1] G.J.Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.
- [2] F. Bossen, B. Bross, S. Karsten, and D. Flynn, "HEVC complexity and implementation analysis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1685–1696, Dec. 2012.
- [3] M. A. Hoque, M. Siekkinen, and J. K. Nurminen, "Energy efficient multimedia streaming to mobile devices a survey," *IEEE Commun. Surv. Tutorials*, vol. 16, no. 1, pp. 579–597, Feb 2014.
- [4] Z. Longhao, R. Trestian, and G.-M. Muntean, "eDOAS: Energy-aware device-oriented adaptive multimedia scheme for wi-fi offload," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Istanbul, Turkey, April 2014, pp. 2916–2921.
- [5] E. Nogues, R. Berrada, M. Pelcat, D. Menard, and E. Raffin, "A DVFS based HEVC decoder for energy-efficient software implementation on embedded processors," in *Proc. IEEE International Conference on Multimedia and Expo (ICME)*, Turin, Italy, July 2015, pp. 1–6.
- [6] Y. He, M. Kunstner, S. Gudumasu, E. S. Ryu, Y. Ye, and X. Xiu, "Power aware HEVC streaming for mobile," in *Proc. IEEE International Conference on Visual Communications and Image Processing*, Kuching, Malaysia, Nov. 2013, pp. 2–6.
- [7] E. Raffin, E. Nogues, W. Hamidouche, S. Tomperi, M. Pelcat, and D. Menard, "Low power HEVC software decoder for mobile devices," *Journal of Real-Time Image Processing*, vol. 12, no. 2, pp. 495–507, 2016.
- [8] T. Mallikarachchi, H. K. Arachchi, D. Talagala, and A. Fernando, "CTU level decoder energy consumption modelling for decoder energy-aware HEVC encoding," in *Proc. IEEE International Conference on Consumer Electronics*, Las Vegas, USA, Jan 2016.
- [9] —, "Decoder energy-aware intra-coded HEVC bit stream generation," in *Proc. IEEE International Conference on Multimedia and Expo*, Seattle, USA, Jul. 2016.
- [10] B. Li, H. Li, L. Li, and J. Zhang, "Lambda-domain rate control algorithm for high efficiency video coding," *IEEE Trans. on Image Processing*, vol. 23, no. 9, pp. 3841–3854, Jul. 2014.
- [11] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, vol. 15, no. 6, pp. 74–90, Nov. 1998.
- [12] JCT-VC, "HEVC reference software - hm-16.0," [https://hevc.hhi.fraunhofer.de/svn/svn\\_HEVCSoftware/tags/HM-16.0/](https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.0/).
- [13] openHEVC, <https://github.com/OpenHEVC/openHEVC>.
- [14] Valgrind Developers, "The valgrind quick start guide," valgrind Documentation.
- [15] G. Bjontegarrd, "Calculation of average PSNR differences between RD-curves," *ITU - Telecommunications Standardization Sector STUDY GROUP 16 Video Coding Experts Group (VCEG)*, 2001.
- [16] E. Nogues, S. Holmbacka, M. Pelcat, D. Menard, and J. Lilius, "Power-aware HEVC decoding with tunable image quality," in *Proc. IEEE workshop on Signal Processing Systems*, Belfast, Noth. Ireland, Oct 2014, pp. 1–6.