**ORGANISATION INTERNATIONALE DE NORMALISATION  
ISO/IEC JTC 1/SC 29/AG 5  
MPEG VISUAL QUALITY ASSESSMENT**

**ISO/IEC JTC 1/SC 29/AG 5 N140**

**Nov. 2024, Kemer, TR**

|  |  |
| --- | --- |
| **Title** | **Report on subjective quality testing of the FGC SEI message** |
| **Source** | **Test coordinators** |
| **Author** | **Philippe de Lagrange, Walt Husak, Mathias Wien** |
| **Status** | **Output document** |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Executive Summary

The film grain characteristics (FGC) SEI message can be embedded in AVC, HEVC, and VVC bitstreams to convey parameters that a decoder can use to synthesize controlled film grain in a post-decoding stage. Such film grain synthesis can be used to recreate a grain that is visually similar to the original, that has been lost due to compression, thus restoring the artistic intent. It can also be used to mask coding defects to some extent.

In this testing effort, the impact of film grain synthesis using parameters provided by the Film Grain Characteristics SEI message for both HEVC and VVC bitstreams has been assessed. In category 1 of the tests, the effect of film grain synthesis on a broad visual quality range has been explored. In category 2, the focus was on the effect of film grain synthesis at high visual quality. The category 1 tests reveal that significant performance improvement can be achieved using FGS technology. For 5 out of 6 test sequences, BD-MOS gains between 1 and 3 on the 11-grade MOS scale have been observed. The tests demonstrate that the overlay of synthetic film grain effectively hides compression artifacts. The results thereby indicate that film grain synthesis seems to generally have a positive effect on the visual quality of the compressed test sequences. The reported tests for category 2 demonstrate that for 5 out of 10 test sequences, the characteristics of the synthesized film grain reach the level of being indistinguishable from the original film grain for the participating viewers at the high rate point of the just noticeable quality difference. For these cases, bitrate savings in the range of 85-96% are reported. Similar gains are reported for both, HEVC and VVC, in both categories. By tendency, gains may be slightly higher for VVC compared to HEVC.

# Introduction

The film grain characteristics (FGC) SEI message can be embedded in AVC, HEVC, and VVC bitstreams to convey parameters that a decoder can use to synthesize controlled film grain in a post-decoding stage. The specification text of the FGC SEI message can found in the AVC, HEVC, and VSEI standards [1][2][3].

Such film grain synthesis can be used to recreate a grain that is visually similar to the original, that has been lost due to compression, thus restoring the artistic intent. It can also be used to mask coding defects to some extent.

This feature has been available for a long time but is currently driving more attention since recent advances in video compression have come with a significant reduction of noise, making film grain preservation more difficult at broadcast or streaming bitrates. Also, recent codecs produce smoother images, where coding artefacts may be more visible; in this context, grain synthesis could help to improve perceived quality by hiding artefacts in smooth areas and increasing perceived sharpness, in addition to restoring the artistic intent when relevant.

In this test effort, the result of VVC and HEVC video compression with and without FGC SEI message and film grain synthesis are compared, in order to evaluate how much rate can be saved using this feature, with film grain characteristics which mimic the original film grain. For this purpose, video content containing different grain types was selected. In this report, results for two test categories are reported, according to the test plan [4]:

1. Assessment of subjective quality improvement provided by film grain synthesis. This assessment evaluates the benefits of film grain synthesis as a tool for improving the visual quality of a compressed video sequence for a broad range of bitrates.
2. Evaluate how much bitrate can be saved by using film grain synthesis instead of natively encoding the grain, while staying at a high visual quality level.

In this report, results for the second category are reported first, since configuration aspects for the first category have been determined on this basis. The testing has been conducted considering both, HEVC and VVC bitstreams in order to demonstrate the suitability of the SEI message across the recent video coding standards.

# Test material

## Test sequences

Table – Test sequences and association of the tested video coding standard.

|  |  |  |  |
| --- | --- | --- | --- |
| **Content type** | **Test sequence & format** | **HEVC / VVC** | **Thumbnail** |
| Scanned film | OldTownCross (from 2004 SVT FairyTale)  3840×2160, 50fps, 500 frames  Kodak Vision 250D 65mm film (negative) | VVC | Ein Bild, das draußen, Himmel, Wolke, Skyline enthält.  Automatisch generierte Beschreibung |
| Scanned film | Two scenes from STeM1 (ASC/DCI)  1920x1080, 24fps, 240 frames  35mm film (negative [TBC]) | HEVC | Ein Bild, das Person, Braut, Kleidung, Hochzeitskleid enthält.  Automatisch generierte Beschreibung |
| Historic scanned film | A scene from 1906 black-and-white “A trip down market street” (Prelinger Archives)  3840x2160, 16fps, 160 frames  35mm film (print) | VVC | Ein Bild, das Weg, Fahrzeug, draußen, Tram enthält.  Automatisch generierte Beschreibung |
| Historic scanned film | A scene from 1963 AmericanChoice outtakes R3 (Prelinger Archives).  3840x2160, 24fps, 240 frames  35mm film (negative) | HEVC | Ein Bild, das Fenster, Gebäude, draußen, Eigentum enthält.  Automatisch generierte Beschreibung |
| Historic scanned film | A scene from 1960 “American Maker” (Prelinger Archives)  1920x1080, 24fps, 240 frames  35mm film (print) | VVC | Ein Bild, das Fahrzeug, Landfahrzeug, Kleidung, draußen enthält.  Automatisch generierte Beschreibung |
| Digital camera noise | DinnerScene2 (from Netflix Chimera)  3840×2160, 60fps, 600 frames  RED Epic Dragon digital camera, sensor noise | HEVC | Ein Bild, das Im Haus, Wand, Person, Kleidung enthält.  Automatisch generierte Beschreibung |
| Postproduction grain | Two 10s clips from Netflix Meridian: from smoker scene, and first car scene  3840×2160, 60fps, 600 frames  Post-production grain | VVC  HEVC | Ein Bild, das Person, Menschliches Gesicht, Wand, Fenster enthält.  Automatisch generierte Beschreibung  Ein Bild, das draußen, Straße, Landfahrzeug, Berg enthält.  Automatisch generierte Beschreibung |
| Postproduction grain | Two scenes from TearsOfSteel, denoised and re-grained:  004 (with and without grain 07) 044 (with and without grain 02)  3840×1714, 24fps, 240 frames | VVC  HEVC | **Ein Bild, das Kleidung, draußen, Person, Baum enthält.  Automatisch generierte Beschreibung**  Ein Bild, das Menschliches Gesicht, Person, Kleidung, Mann enthält.  Automatisch generierte Beschreibung |

## Grain parameters

The film grain parameters conveyed by FGC SEI messages were determined manually by following the process described in JVET-AG0324 [5]. This involved using a tool that displays the parameters as a graph that can be edited interactively (e.g. using the mouse to drag lines representing film grain amplitude, intensity interval boundaries, or frequency cutoffs); as soon as the user makes a change, the tool generates a new film grain configuration file and executes film grain synthesis so that the impact of the changes can be judged immediately.

A screenshot of a computer

Description automatically generated

Figure 1 –Main window of film grain parameter adjustment tool

The parameters were adjusted to visually match original picture, after applying film grain synthesis on a VTM-encoded picture that has grain removed or mostly removed. A variable effort was spent for the various test sequences, only exercising frequency filtering with shared horizontal/vertical frequency cutoff because of the current limitations of the adjustment tool, though the synthesis software has broader support (more parameters, and auto-regressive mode). This can be seen as a lower bound on film grain characteristics fidelity and was deemed sufficient for the purpose of the test.

Parameters exercised include either luma-only or colored grain, various light intensity intervals, and variable gain and frequency cutoff depending on intensity interval.

The parameters are provided as an attachment to this document. They are mostly identical to those determined in JVET-AI0138 [6], except for DinnerScene2 where separate adjustments were made for the second part of the clip (after the scene cut). Each configuration file represents data conveyed by one FGC SEI message, which can be persistent for several pictures and repeated several times for a given sequence.

The bitrate impact of FGC SEI messages for each sequence can be evaluated from the following table, where the FGC SEI message size was measured by actual insertion in the bitstreams using the SEIFilmGrainApp program available in the VTM:

Table – Filmgrain characteristics information

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Test sequence** | **Configuration file** | **FGC SEI bits** | **Repetitions** | **Total bits** |
| AmericanChoice | americanchoice.cfg | 576 | 8 | 4608 |
| AmericanMaker | americanmaker.cfg | 696 | 8 | 5568 |
| DinnerScene2 | DinnerScene.cfg | 720 | 10 | 7200 |
| DinnerScene2 | DinnerScene2-a.cfg | 720 | 6 | 8200 |
| DinnerScene2-b.cfg | 776 | 5 |
| MagicHour | MagicHour2.cfg | 760 | 8 | 7600 |
| MarketStreet1 | marketstreet.cfg | 408 | 5 | 2040 |
| MeridianIntoCar1 | meridian-intocar-1a.cfg | 712 | 7 | 8344 |
| meridian-intocar-1b.cfg | 840 | 4 |
| MeridianSmoker1 | meridian-smoker-1.cfg | 912 | 10 | 9120 |
| OldTownCross | OldTownCross.cfg | 504 | 8 | 4032 |
| TearsOfSteel-004 | tearsofsteel-07.cfg | 640 | 8 | 5120 |
| TearsOfSteel-044 | tearsofsteel-02-new.cfg | 640 | 8 | 5120 |

## Grain synthesis

Film grain synthesis was applied using the VFGS software, version 2.1, available in https://vcgit.hhi.fraunhofer.de/jvet-ahg-fgt/vfgs (software repository for the JVET film grain ad-hoc group), as it is deemed representative of a practical hardware implementation supporting a wider range of parameters in the FGC SEI message than what is currently available in the VTM (e.g. limited to frequency-filtering mode). It is planned to merge this software into the VTM as soon as possible.

In the Annex, the application of film grain parameters with the VFGS software is explained.

# Comparison of film grain coding and native encoding at high visual quality

## Test setup

### Logistics

The tests were conducted with video coding experts at the 35th JVET meeting in Sapporo, JP.

The setup at the meeting site included a PC with a Decklink video board for HDMI connection and SSD drives capable of stable playout of the raw YUV data at the required frame rate.

Table – Test setup.

| **Test Site** | **On-site** |
| --- | --- |
| **Display, size, connection  (resolution setting)** | 1x Sony 65” LCD, 1x Sharp 65” LCD HDMI (3840×2160), 10bit input |
| **Viewing distance** | 3 viewers sitting at 1.5H for each display |
| **Viewing angle** | ±75°, 90° (at screen center) |
| **Total number of viewers** | 23 MPEG experts |

Participants confirmed visual acuity and normal color vision.

The setup was chosen with three viewers sitting in front of each display. The viewers entered their scores into the test database after the viewing session.

Visual assessment was performed on video sequences in YUV 4:2:0 10 bit planar format. The test sequences were selected to cover a range of different source and application types. Both VTM and HM bitstreams are considered. A pre-viewing effort was made to identify QP values around the estimated point of transparency. The resulting selection is as follows:

* **Film material**: MarketStreet, AmericanChoice, AmericanMaker, OldTownCross, MagicHour
* **Digitally captured and processed video**: MeridianSmoker1, MeridianIntoCar1, DinnerScene2
* **“clean” video with known synthetic film grain**: TearsOfSteel-004, TearsOfSteel-044

### Test sequences and quantizer settings

|  |  |  |  |
| --- | --- | --- | --- |
| **Test sequence** | **Test model** | **QP no FGS** | **QP FGS** |
| AmericanChoice | HM-18.0 | 16, 18, 20, 22 | 28, 30, 32, 34 |
| AmericanMaker | VTM-23.1 | 13, 15, 17, 19 | 26, 28, 30, 32 |
| DinnerScene2 | HM-18.0 | 14, 16, 18, 20 | 20, 22, 24, 26 |
| MagicHour | HM-18.0 | 12, 14, 16, 18 | 20, 23, 26, 29 |
| MarketStreet | VTM-23.1 | 20, 22, 24, 26 | 29, 32, 35, 38 |
| MeridianIntoCar1 | HM-18.0 | 16, 18, 20, 22 | 22, 24, 26, 28 |
| MeridianSmoker1 | VTM-23.1 | 18, 20, 22, 24 | 22, 24, 26, 29 |
| OldTownCross | VTM-17.0 | 14, 16, 18, 20 | 22, 24, 26, 28 |
| TearsOfSteel004 | VTM-23.1 | 16, 18, 20, 22 | 24, 26, 28, 30 |
| TearsOfSteel044 | HM-18.0 | 12, 14, 16, 18 | 22, 24, 26, 28 |

### Test design

#### Protocol and session setup

The test was conducted following the Degradation Category Rating (DCR) / Double Stimulus Impairment Scale (DSIS) method of ITU-T P.910 / ITU-R BT.500 [8][9] with an adapted scale. Since the viewing task was to identify differences at a very high-quality range, a three-grade scale was employed:

* 0: no difference visible
* 1: small differences visible
* 2: visually different

For evaluation, three test sessions with 31 basic test cells (BTCs) each were presented to the viewers. The duration of each test session was about 16 min. The test session included a stabilization phase of three BTCs at the beginning. The test sessions were preceded by an explanatory introduction and a training session for the type of content to be viewed.

Each BTC was structured as follows (PVS denotes the processed video sequence under evaluation):

**“Original” (1s) – [uncompressed seq.] (10s) – “PVS” (1sec) – [PVS] (10sec) – “Vote <N>” (5s).**

The indicated text was displayed on a mid-gray background. All sequences were played out at their original resolution and frame rates.

#### Viewer information and training

The viewers were explained the viewing task, the organization of the session and the meaning of the grading scale. All questions of the viewers regarding the test session, procedure, etc., were answered. For the training, a test session with 10 BTCs was employed, covering examples of the impairment range.

## Subjective results and analysis

### Data processing

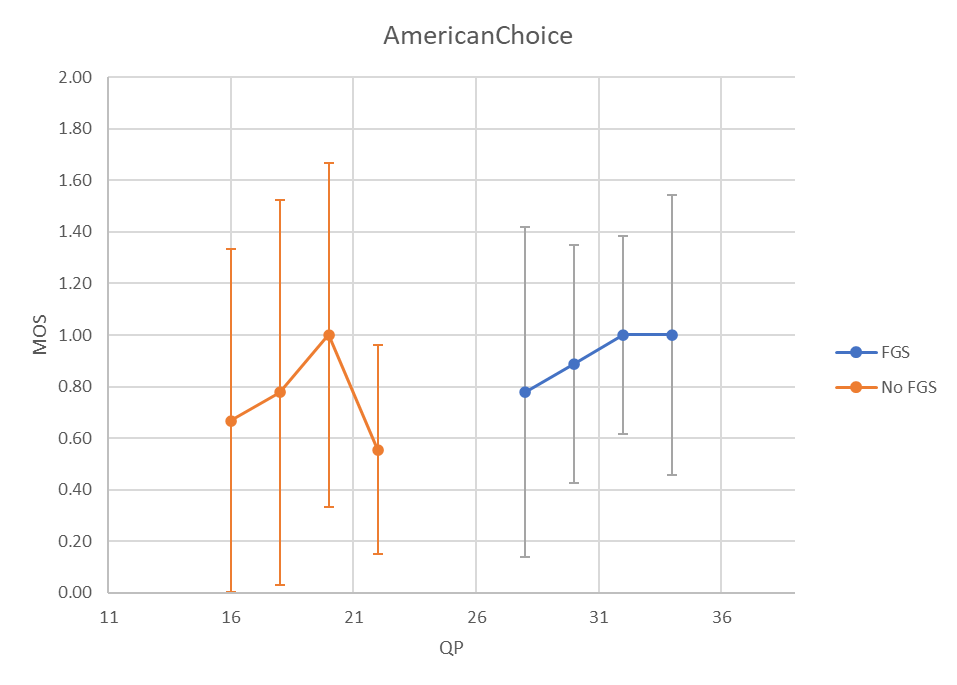
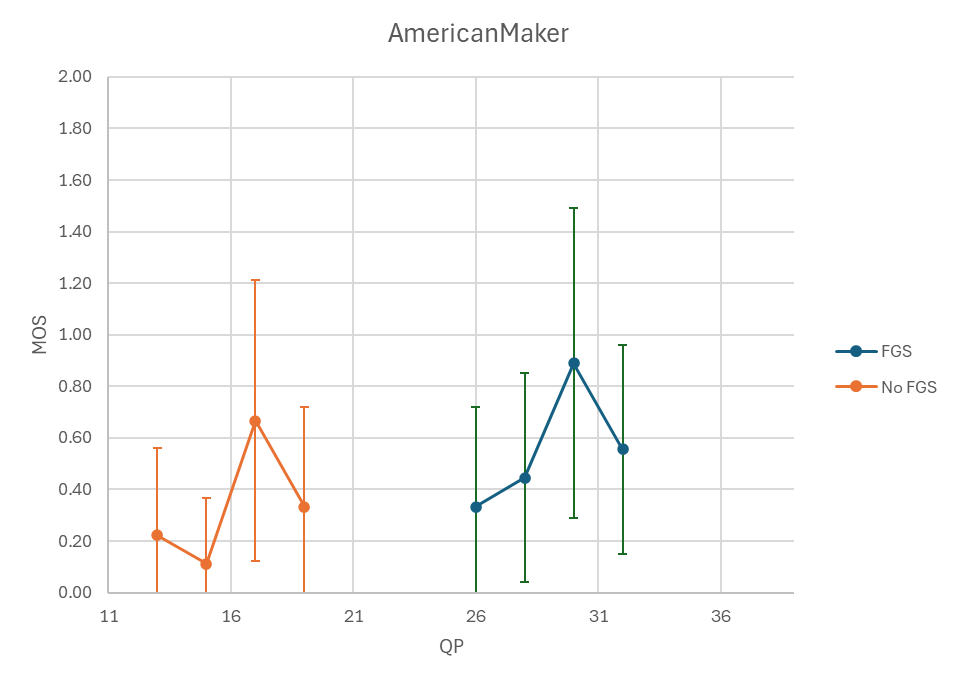
The viewing task is considered to be challenging for the viewers. This is specifically due to the narrow quality range that was assessed.

The following steps of data processing were applied:

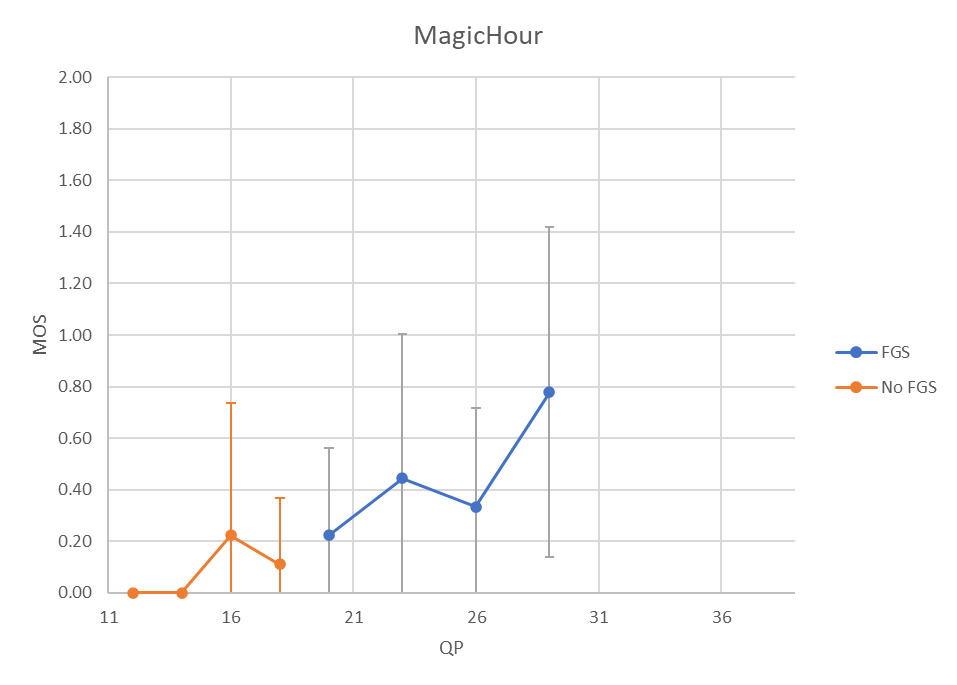
* Trapping sequence: A comparison of the uncompressed vs. the same uncompressed sequence was included. No viewer scored these uncompressed sequences higher than 1, which was considered to be acceptable.
* Checking the ability to see differences: Based on the definition of Opinion Score OS=0 indicating no visual difference, viewers with of a total sum of OS lower or equal to 10were not regarded in the further evaluation (assessment excluding the stabilization sequences, i.e. out of 84 scores) This applied to 10 viewers.
* Checking the progression consistency of scores: The check is based on the established assumption that an increase of the QP in VTM and HM implies a (potentially small) visual quality drop. On this basis, an OS delta check was applied over the four QPs evaluated for each test point. The number of cases where a viewer would score the PVS with a lower QP worse than the PVS with the higher QP was counted. The viewers with more than 15 cases of such inversion were not regarded in the further evaluation. This applied to three viewers.

### Subjective results

The resulting MOS values on the 3-grade scale are plotted over the QP for the corresponding bitstream. The ±95% confidence intervals (CIs) for the MOS values are indicated. Additionally, the histograms of the OS for all PVS are provided.

Ein Bild, das Text, Diagramm, Reihe, Zahl enthält.

Automatisch generierte Beschreibung 

Ein Bild, das Text, Diagramm, Reihe, parallel enthält.

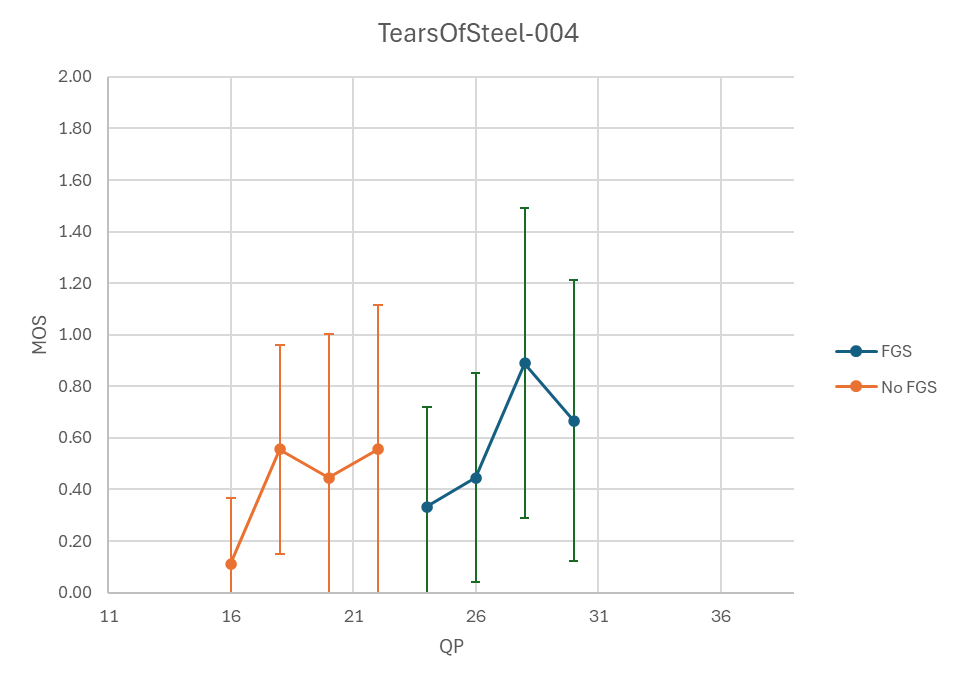
Automatisch generierte Beschreibung Ein Bild, das Text, Diagramm, Reihe, Zahl enthält.

Automatisch generierte Beschreibung

Ein Bild, das Text, Diagramm, Reihe, Zahl enthält.

Automatisch generierte Beschreibung Ein Bild, das Text, Diagramm, Reihe, Zahl enthält.

Automatisch generierte Beschreibung

 Ein Bild, das Text, Diagramm, Reihe, Zahl enthält.

Automatisch generierte Beschreibung

Figure MOS over QP plots and scoring statistics for the test sequences under evaluation.

### Evaluation of results

In order to summarize the outcome, the data points were assessed according to the following criteria:

* MOS values at or below 0.5 are considered to fall below the threshold of just noticeable differences.
* PVSs with a MOS over 0.5 and a CI value not ranging over the MOS=0 line are considered significant.
* The PVS with the lowest QP meeting the significance condition is reported to be the point where a visual difference might be observed.

Based on these criteria, the results of the experiment can be summarized in the following table. When possible, the QP of visual difference is reported.

When a QP is reported with and without grain synthesis, the rate difference of those test points is reported as . Those numbers represent how much bit rate is saved when not encoding the film grain as pixels, but using grain synthesis instead, at a video quality close to visual transparency; they should be taken as indicative, since they were obtained with the VTM and HM encoders using default settings, and different numbers may be obtained with different encoders or different settings.

Table – JND QP points and corresponding estimated rate savings per sequence.

|  |  |  |  |
| --- | --- | --- | --- |
| **Test sequence** | **No film grain synthesis** | **With film grain synthesis** | **Estimated rate savings** |
| AmericanChoice | 18 | all significant | (not determined) (HM) |
| DinnerScene2 | 16 | all significant | (not determined) (HM) |
| MagicHour | no significance | 29 | (not determined) (HM) |
| MeridianIntoCar1 | 18 | 24 | 95% (HM) |
| TearsOfSteel-044 | 18 | all significant | (not determined) (HM) |
| AmericanMaker | 17 | 30 | 94% (VTM) |
| MarketStreet1 | 26 | 32 | 85% (VTM) |
| MeridianSmoker1 | 22 | 26 | 90% (VTM) |
| OldTownCross | 20 | all significant | (not determined) (VTM) |
| TearsOfSteel-004 | 18 | 28 | 96% (VTM) |

It is noted that for multiple sequences, all PVS under test were rated visually significantly different by the viewers. This may be due to two reasons: a) the viewers observed a (potentially slight) difference between the synthesized and the original film grain, or b) the viewers observed compression artifacts for all evaluated test points. This aspect is supposed to be studied in the remaining category 3 of the subjective quality testing of the FGC SEI message.

# Assessment of film grain synthesis as a tool for improving visual quality over a wider range of bitrates

Based on the results for determining the JND QP value reported in Section 3, the QP points to be tested in the category 1 tests were determined.

## Test setup

### Logistics

The tests were conducted with video coding experts at the 36th JVET meeting in Kemer, TR.

The setup at the meeting site included a PC with a Decklink video board for HDMI connection and SSD drives capable of stable playout of the raw YUV data at the required frame rate.

Table – Test setup.

| **Test Site** | **On-site** |
| --- | --- |
| **Display, size, connection  (resolution setting)** | TCL 85C855, 2×TCL 75C855 (all miniLED); VESTEL 55” OLED, HDMI (3840×2160), 10bit input |
| **Viewing distance** | Viewers sitting at 1.5H (4 viewers for the 85” display, 3 viewers for the 75” displays, 2 viewers for the 55” display) |
| **Viewing angle** | ±75°, 90° (at screen center) |
| **Total number of viewers** | 25 (8 female, 17 male) |

Participants confirmed visual acuity and normal colour vision.

The display settings were browsed, and any enhancement features were switched off. For the VESTEL 55” display, an internal frame rate up-conversion was switched on in the test. No disabling option was found.

### Test sequences and quantizer settings

For evaluation, a subset of test sequences and quantizer settings listed in JVET-AI2022 [1] and JVET‑AI0138 [6] were tested. Six UHD sequences were tested, including the test sequences DinnerScene2, MarketStreet1, MeridianIntoCar1, MeridianSmoker1, OldTownCross and TearsOfSteel-044 which have been prepared in JVET‑AI0138 [6].

Two coding configurations were tested:

* NonFGS with the sequences being encoded and not enhanced with an FGS information
* and FGS with the encoded sequences being enhanced by SEI FGS information as described in JVET‑AI0138 [6].

For TearsOfSteel-044, which exists in two version, clean original and original enhanced with synthetic grain, the NonFGS version is encoded from the grainy original, and the FGS version is encoded from the clean original. For all other sequences the original only exits in grainy version, which is used for both encodings.

In addition to two tests, a “just noticeable difference (JND)” QP point of the NonFGS version has been included in the tests, with the JND QP established in the test JVET‑AI0337 [7]. It represents an encoding which the viewers in the previous test established as just distinguishable from the original. The QP values evaluated in the expert viewing are shown in the table below.

Table – Selected test sequences and QP settings for the viewing.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sequence** | **Codec** | **QP JND** | **QP no FGS** | **QP FGS** |
| DinnerScene2 | HM-18.0 | 16 | 22, 26, 30, 34 | 22, 26, 30, 34 |
| MeridianIntoCar1 | HM-18.0 | 18 | 24, 28, 32, 36 | 24, 28, 32, 36 |
| TearsOfSteel044 | HM-18.0 | 18 | 24, 28, 32, 36 | 24, 28, 32, 36 |
| MarketStreet | VTM-23.1 | 26 | 30, 34, 38, 42 | 30, 34, 38, 42 |
| MeridianSmoker1 | VTM-23.1 | 20 | 28, 32, 36, 40 | 28, 32, 36, 40 |
| OldTownCross | VTM-17.0 | 20 | 29, 33, 37, 41 | 29, 33, 37, 41 |

### Test method and test design

The assessment was performed with the reconstructed UHD output. Two DCR test session with 37 BTCs were presented to the viewers. The duration of each test session was about 19 min.

#### Degradation category rating

##### Session setup

The test sequences were evaluated using an 11-grade scale shown in Figure 3 below.

Ein Bild, das Text, Screenshot, Schrift, Zahl enthält.

Automatisch generierte Beschreibung

Figure 3: Meaning of the used 11 grade numerical scale

The test sessions included 37 basic test cells (BTCs) showing the original uncompressed video sequence and one processed video sequence. Each BTC was structured as followed (PVS denotes the processed video sequence under evaluation):

**“Original” (1sec) – [original sequence] (10sec) – “A” (1sec) – [PVS] (10sec) – “Vote <N>” (5sec)**

The session included a stabilization phase of three BTCs, and trapping BTCs where original full resolution test sequences were to be scored in the PVS.

##### Viewer information and training

The viewers were explained the viewing task, the organization of the session and the meaning of the grading scale. The instructions focused on the DCR method. They did not include specific guidance with respect to film grain being tested. The viewers were organized into two groups of 12 and 13 participants, respectively. Training, viewing of session 1 and viewing of session 2 were done in an interleaved fashion, where one group performed the viewing tasks at the four displays while the other group had a break. During the viewing sessions, scoring was done on paper sheets. The breaks were used to enter the acquired scores into the test database.

For training, the participants were shown a test session of 12 BTCs including representative examples of the impairments shown in the tests.

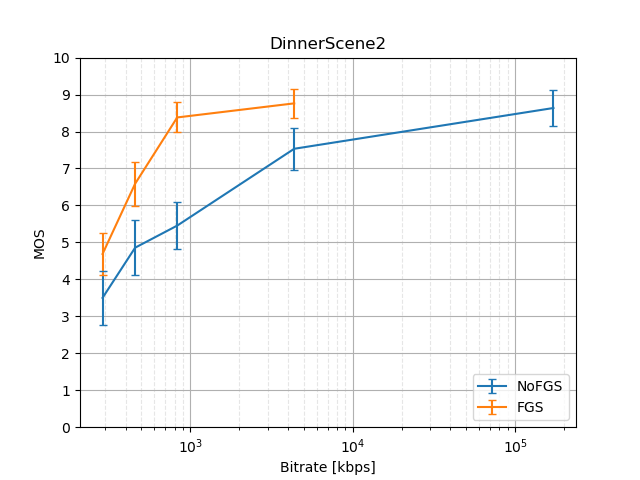
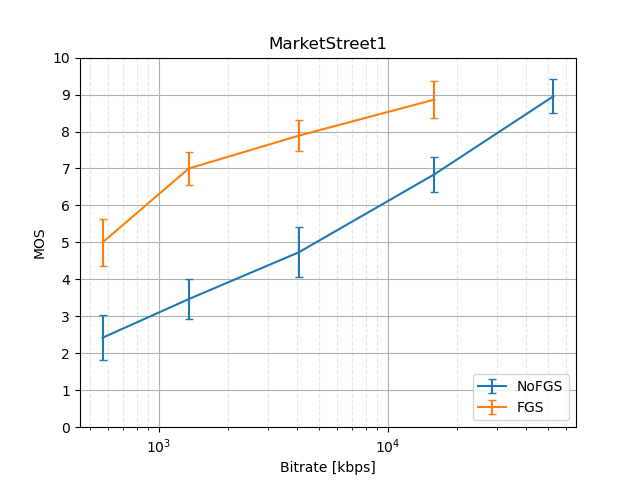
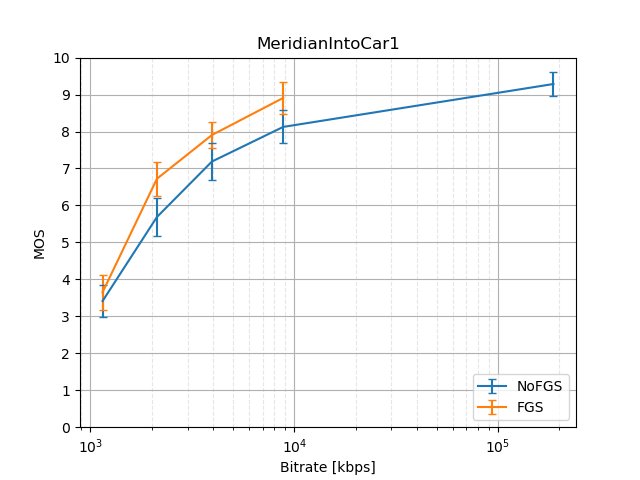
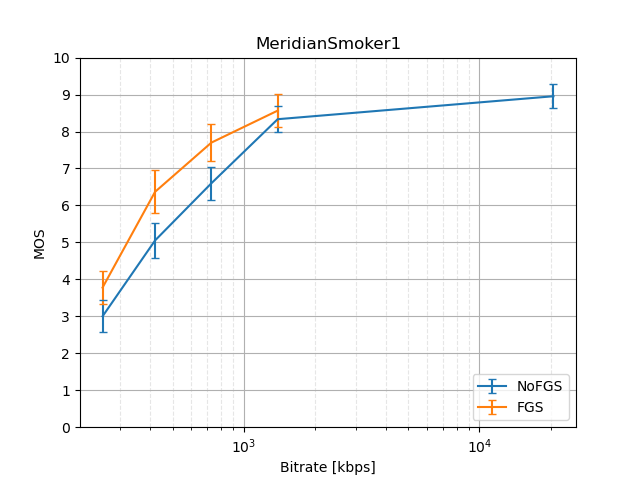
## Subjective results and analysis

### Data processing

It was noticed that the Pearson correlation coefficient of the viewers scores for session 2 was significantly higher than for session 1 (mean increasing from 80% to 88%). It was decided to only remove viewers with a Pearson correlation of below 70% relative to the MOS. Three viewers with an otherwise a high correlation entered an opinion score of 7 for the trapping sequence (original vs. original) in the first session. It was decided to keep them in the set. As a last step, the non-normalized -scores of the viewer scores were compared to a threshold of 2, where is the score of viewer for PVS , is the mean opinion score for that PVS for the viewers. Based on this threshold, 19.2% of the scores were removed from the dataset.

### Subjective results

The measured MOS values of the reconstructed video on the 11-grade scale are plotted over the bit rate of the corresponding bitstream. The ±95% confidence intervals for the MOS values are indicated. The plots show the MOS-over-bitrate plots for the cases with (FGS) and without (NoFGS), respectively. The highest rate point of the NoFGS curve corresponds to the JND QP point reported in Table 6.

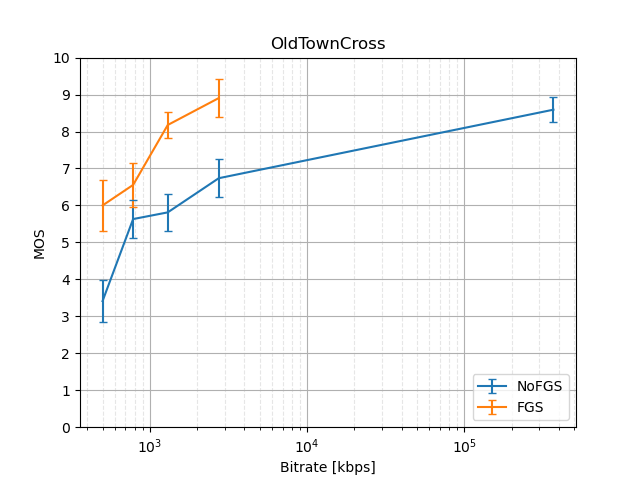
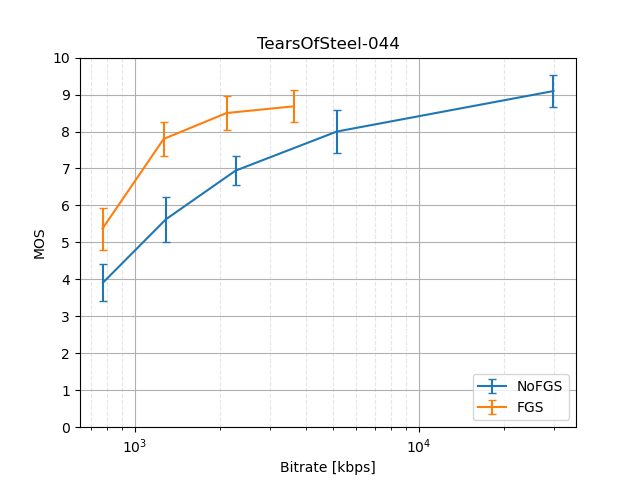
 

Figure 4: MOS-over-rate plots for the DCR experiment

Due to the fact that the FGC SEI message only has minor impact on the bitrate, the NoFGS and FGS bitstreams share very similar rate points. Only for the TearsOfSteel-044 sequence, the rate point diverge. This is due to the fact that in this case, the encoded sequences differ: in the noFGS case, the re-grained sequence is encoded while in the FGS case, the denoised version is encoded with the FGC SEI message to control the FGS at the decoder side.

Table 7 below reports the BD-MOS values which indicate the achievable increase in terms of MOS (on the 11-grade scale) when applying FGS compared to the noFGS case.

Table – Bjontegaard Delta MOS values.

|  |  |  |
| --- | --- | --- |
| **Sequence** | **Codec** | **BD-MOS** |
| DinnerScene2 | HM | 2.12 |
| MeridianIntoCar1 | HM | n/a |
| TearsOfSteel044 | HM | 1.77 |
| MarketStreet | VTM | 3.04 |
| MeridianSmoker1 | VTM | 1.01 |
| OldTownCross | VTM | 1.95 |

Since the CI for the MeridianIntoCar1 overlap, a BD-MOS is not reported for this sequence. For the other sequences, BD-MOS values between about 1 and 3 are observed.

# Discussion

The impact of film grain synthesis using parameters provided by the Film Grain Characteristics SEI message for both HEVC and VVC bitstreams has been assessed. In category 1 of the tests, the effect of film grain synthesis on a broad visual quality range has been explored. In category 2, the focus was on the effect of film grain synthesis at high visual quality. The category 1 tests reveal that significant performance improvement can be achieved using FGS technology. For 5 out of 6 test sequences, BD-MOS gains between 1 and 3 on the 11-grade MOS scale have been observed. The tests demonstrate that the overlay of synthetic film grain effectively hides compression artifacts. The results thereby indicate that film grain synthesis seems to generally have a positive effect on the visual quality of the compressed test sequences. The reported tests for category 2 demonstrate that for 5 out of 10 test sequences, the characteristics of the synthesized film grain reach the level of being indistinguishable from the original film grain for the participating viewers at the high rate point of the just noticeable quality difference. For these cases, bitrate savings in the range of 85-96% are reported. Similar gains are reported for both, HEVC and VVC, in both categories. By tendency, gains may be slightly higher for VVC compared to HEVC.

# References

1. Rec. ITU-T H.264 | ISO/IEC 14496-10, “Information technology — Coding of audio-visual objects — Part 10: Advanced video coding”
2. Rec. ITU-T H.265 | ISO/IEC 23008-2, “Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding”
3. Rec. ITU-T H.274 | ISO/IEC 23002-7, “Versatile supplemental enhancement information messages for coded video bitstreams”
4. P. de Lagrange, W. Husak, M. Wien, “Draft plan for subjective quality testing of the FGC SEI message,” Doc. JVET-AI2022, 35th meeting, Sapporo, Jul. 2024.
5. P. de Lagrange, “AHG4/AHG13/AHG3: film grain tuning tool tutorial,” Doc. JVET-AG0324, 33rd meeting, online, Jan. 2024.
6. P. de Lagrange, “AHG4: preparation of test material for film grain visual testing,” Doc. JVET-AI0138, 35th meeting, Sapporo, Jul. 2024.
7. M. Wien, P. de Lagrange, A. Wieckowski, “Preliminary results of expert viewing for FGC subjective test category 2,” Doc. JVET-AI0337, 35th meeting, Sapporo, Jul. 2024.
8. ITU-T Rec. P.910, “Subjective video quality assessment methods for multimedia applications,” Oct. 2023.
9. ITU-R Rec. BT.500, “Methodologies for the subjective assessment of the quality of television images,” May 2023.

# Annex 1: Application of film grain synthesis with the VFGS software

Film grain synthesis was applied using the VFGS software, version 2.1, available in https://vcgit.hhi.fraunhofer.de/jvet-ahg-fgt/vfgs/-/tree/VFGS-2.1?ref\_type=tags.

This software replicates what can be done in an AVC, HEVC, or VVC decoder when an FGC SEI message is received. The film grain synthesis process follows the semantics of the FGC SEI message with some optimizations and parameter restrictions intended for practical hardware implementation; more information can be found in the VFGS readme file and in the ISO 23007-2 technical report. The software takes as inputs a YUV file and FGC SEI message contents in the form of one or several configuration files (and optionally the image index where each applies), and outputs a YUV file.

The command line used for the various test sequences were the following for Sapporo tests:

vfgs -w 3840 -h 2160 -c americanchoice.cfg AmericanChoice\_[QP].yuv AmericanChoice\_[QP]\_fgs.yuv

vfgs -w 1920 -h 1080 -c americanmaker.cfg AmericanMaker\_[QP].yuv AmericanMaker\_[QP]\_fgs.yuv

vfgs -w 3840 -h 2160 -c DinnerScene.cfg DinnerScene2\_[QP].yuv DinnerScene2\_[QP]\_fgs.yuv

vfgs -w 1920 -h 1080 -c MagicHour2.cfg MagicHour\_[QP].yuv MagicHour\_[QP]\_fgs.yuv

vfgs -w 3840 -h 2160 -c marketstreet.cfg MarketStreet1\_[QP].yuv MarketStreet1\_[QP]\_fgs.yuv

vfgs -w 3840 -h 2160 -c 0:meridian-intocar-1a.cfg -c 392:meridian-intocar-1b.cfg MeridianIntoCar1\_[QP].yuv MeridianIntoCar1\_[QP]\_fgs.yuv

vfgs -w 3840 -h 2160 -c meridian-smoker-1.cfg MeridianSmoker1\_[QP].yuv MeridianSmoker1\_[QP]\_fgs.yuv

vfgs -w 3840 -h 2160 -c OldTownCross.cfg OldTownCross\_[QP].yuv OldTownCross\_[QP]\_fgs.yuv

vfgs -w 3840 -h 1714 -c tearsofsteel-07.cfg TearsOfSteel-004\_[QP].yuv TearsOfSteel-004\_[QP]\_fgs.yuv

vfgs -w 3840 -h 1714 -c tearsofsteel-02-new.cfg TearsOfSteel-044\_[QP].yuv TearsOfSteel-044\_[QP]\_fgs.yuv

The command line for DinnerScene2 was modified as follows for Kemer tests:

vfgs -w 3840 -h 2160 -c 0:DinnerScene2-a.cfg -c 360:DinnerScene2-b.cfg DinnerScene2\_[QP].yuv DinnerScene2\_[QP]\_fgs.yuv

# Annex 2 Category 1 MOS data

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | NoFGS | | | FGS | | |
| Testmodel | Sequence | QP | Rate | MOS | CI | Rate | MOS | CI |
| HM | DinnerScene2 | 22 | 4364.10 | 7.53 | 0.57 | 4364.92 | 8.76 | 0.39 |
| 26 | 830.45 | 5.45 | 0.63 | 831.27 | 8.38 | 0.40 |
| 30 | 457.18 | 4.86 | 0.75 | 458.00 | 6.59 | 0.60 |
| 34 | 287.99 | 3.50 | 0.73 | 288.81 | 4.69 | 0.56 |
| MeridianIntoCar1 | 24 | 8834.42 | 8.13 | 0.45 | 8835.26 | 8.91 | 0.44 |
| 28 | 3946.98 | 7.19 | 0.50 | 3947.81 | 7.90 | 0.35 |
| 32 | 2121.24 | 5.69 | 0.52 | 2122.07 | 6.72 | 0.46 |
| 36 | 1147.41 | 3.41 | 0.42 | 1148.24 | 3.65 | 0.48 |
| TearsOfSteel-044 | 24 | 5144.45 | 8.00 | 0.58 | 3629.39 | 8.68 | 0.43 |
| 28 | 2262.59 | 6.94 | 0.40 | 2104.18 | 8.50 | 0.46 |
| 32 | 1287.21 | 5.63 | 0.60 | 1264.71 | 7.80 | 0.46 |
| 36 | 772.51 | 3.91 | 0.50 | 770.76 | 5.38 | 0.57 |
| VTM | MarketStreet1 | 30 | 15966.84 | 6.84 | 0.48 | 15967.04 | 8.86 | 0.49 |
| 34 | 4088.28 | 4.73 | 0.67 | 4088.48 | 7.89 | 0.43 |
| 38 | 1351.91 | 3.47 | 0.55 | 1352.11 | 7.00 | 0.45 |
| 42 | 567.99 | 2.42 | 0.60 | 568.19 | 5.00 | 0.64 |
| MeridianSmoker1 | 28 | 1394.77 | 8.33 | 0.35 | 1395.68 | 8.57 | 0.44 |
| 32 | 730.33 | 6.60 | 0.45 | 731.24 | 7.70 | 0.50 |
| 36 | 421.50 | 5.05 | 0.48 | 422.41 | 6.37 | 0.58 |
| 40 | 252.46 | 3.00 | 0.44 | 253.37 | 3.78 | 0.45 |
| OldTownCross | 29 | 2744.21 | 6.74 | 0.52 | 2744.62 | 8.90 | 0.51 |
| 33 | 1300.03 | 5.81 | 0.50 | 1300.44 | 8.18 | 0.34 |
| 37 | 782.83 | 5.63 | 0.50 | 783.23 | 6.56 | 0.59 |
| 41 | 500.19 | 3.41 | 0.57 | 500.59 | 6.00 | 0.69 |

# Annex 3 Category 2 MOS data

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | NoFGS | | | FGS | | |
| Testmodel | Sequence | QP | MOS | CI | QP | MOS | CI |
| HM | AmericanChoice | 16 | 0.67 | 0.67 | 28 | 0.78 | 0.64 |
| 18 | 0.78 | 0.75 | 30 | 0.89 | 0.46 |
| 20 | 1.00 | 0.67 | 32 | 1.00 | 0.38 |
| 22 | 0.56 | 0.41 | 34 | 1.00 | 0.54 |
| DinnerScene2 | 14 | 0.44 | 0.56 | 20 | 1.00 | 0.67 |
| 16 | 0.56 | 0.41 | 22 | 1.11 | 0.60 |
| 18 | 0.78 | 0.64 | 24 | 1.00 | 0.38 |
| 20 | 1.22 | 0.51 | 26 | 1.33 | 0.38 |
| MagicHour | 12 | 0.00 | 0.00 | 20 | 0.22 | 0.34 |
| 14 | 0.00 | 0.00 | 23 | 0.44 | 0.56 |
| 16 | 0.22 | 0.51 | 26 | 0.33 | 0.38 |
| 18 | 0.11 | 0.26 | 29 | 0.78 | 0.64 |
| MeridianIntoCar1 | 16 | 0.11 | 0.26 | 22 | 0.22 | 0.34 |
| 18 | 0.56 | 0.41 | 24 | 1.00 | 0.54 |
| 20 | 0.22 | 0.34 | 26 | 0.78 | 0.64 |
| 22 | 0.44 | 0.41 | 28 | 1.11 | 0.46 |
| TearsOfSteel-044 | 12 | 0.56 | 0.56 | 22 | 0.67 | 0.54 |
| 14 | 0.33 | 0.38 | 24 | 0.56 | 0.41 |
| 16 | 0.33 | 0.38 | 26 | 0.78 | 0.34 |
| 18 | 0.56 | 0.41 | 28 | 1.11 | 0.71 |
| VTM | AmericanMaker | 13 | 0.22 | 0.34 | 26 | 0.33 | 0.38 |
| 15 | 0.11 | 0.26 | 28 | 0.44 | 0.41 |
| 17 | 0.67 | 0.54 | 30 | 0.89 | 0.60 |
| 19 | 0.33 | 0.38 | 32 | 0.56 | 0.41 |
| MarketStreet1 | 20 | 0.11 | 0.26 | 29 | 0.56 | 0.56 |
| 22 | 0.11 | 0.26 | 32 | 1.00 | 0.67 |
| 24 | 0.44 | 0.56 | 35 | 0.67 | 0.54 |
| 26 | 0.56 | 0.41 | 38 | 1.67 | 0.38 |
| MeridianSmoker1 | 16 | 0.22 | 0.34 | 22 | 0.44 | 0.41 |
| 18 | 0.44 | 0.41 | 24 | 0.33 | 0.38 |
| 20 | 0.56 | 0.56 | 26 | 0.56 | 0.41 |
| 22 | 0.78 | 0.51 | 29 | 0.44 | 0.41 |
| OldTownCross | 14 | 0.22 | 0.34 | 22 | 0.78 | 0.51 |
| 16 | 0.56 | 0.56 | 24 | 0.78 | 0.64 |
| 18 | 0.56 | 0.56 | 26 | 0.89 | 0.46 |
| 20 | 0.56 | 0.41 | 28 | 0.78 | 0.64 |
| TearsOfSteel-004 | 16 | 0.11 | 0.26 | 24 | 0.33 | 0.38 |
| 18 | 0.56 | 0.41 | 26 | 0.44 | 0.41 |
| 20 | 0.44 | 0.56 | 28 | 0.89 | 0.60 |
| 22 | 0.56 | 0.56 | 30 | 0.67 | 0.54 |