

# The SVT High Definition Multi Format Test Set

## 1. Introduction

This Document contains information about the *SVT High Definition Multi Format Test Set* – picked from the demanding, but not unduly so, multi-genre TV-program ‘Fairytale’ (by SVT) mastered in 3840x2160p/50.

All sequences from this set may be distributed freely (complete or in parts) as long as the copyright and the restrictions of the use of the sequences are not violated.

This document has to be distributed with any part of the test set that is distributed.

It is not allowed to charge a fee for the distribution of the sequences. Anybody receiving material from this data set is kindly asked to further distribute these sequences to other parties to reach a super distribution effect.

All sequences from this set were filmed in 50 fps with professional, high end 65mm film equipment by SVT in October 2004. The utmost care has been taken when converting the film into digital data, the same applies for all conversion steps applied afterwards.

The data itself comes in various different resolutions which are widely used – all in 50 Hz motion portrayal. Lower resolutions were gained by filtering the 2160p/50 master. At all resolutions the bit depth is full interval 16 bits per (RGB) colour plane to obtain the very high quality of the original shots.

Details concerning shooting of the sequences and post-processing steps can be found in the section ‘Technical Information’.

If you have any questions concerning these test sequences, please contact:

Lars Haglund, [lars.haglund@svt.se](mailto:lars.haglund@svt.se)

## 2. Copyright and Restrictions of Use

Individuals and organizations extracting sequences from this archive agree that the sequences and all intellectual property rights therein remain the property of Sveriges Television AB (SVT), Sweden. These sequences may only be used for the purpose of developing, testing and presenting technology standards. SVT makes no warranties with respect to the materials and expressly disclaim any warranties regarding their fitness for any purpose.

## 3. Technical Information

Production details	
<i>Time of original on-location production</i>	October 2004. (Post production during 2005).
<i>Producer</i>	Sveriges Television AB (SVT), SVT Technology/R&D. Lars Haglund, <a href="mailto:lars.haglund@svt.se">lars.haglund@svt.se</a>
<i>Camera</i>	ARRI ArriFlex 765 System for 65mm, 5 perf, film
<i>Filming Speed</i>	50 fps (50-100 fps for slow-motion)
<i>Mirror shutter</i>	180 degrees resulting in 1/100 second exposure time
<i>Lenses</i>	Zeiss/ArriFlex prime lenses (30-700mm)
<i>Film Stocks</i>	Exterior: Kodak Vision, 250D

	Interior: Kodak Vision2, 500T
<b>Transfer<sup>1</sup></b>	
<i>Scanner</i>	FilmLight – 65mm NorthLight film scanner
<b>Post processing</b>	
<i>Post-processing software</i>	Apple Shake
<i>Bit depth per colour plane</i>	Full interval 16 bit (linear), no head room or foot room <sup>2</sup>
<i>Transfer and chromaticity parameters</i>	ITU-R BT.709 <sup>3 4</sup>
<i>Filtering for down-sampling</i>	Sinc filter (as in Apple Shake)

### 3.1. Output Formats

For all resolutions:

<i>Transfer and chromaticity parameters</i>	ITU-R BT.709, Gamma corrected values
<i>Bit depth per colour plane</i>	Full interval 16 bit, no head room or foot room
<i>Data format</i>	‘sgi16’ (interleaved RGB), one file per frame Details to this format can be found at <a href="ftp://ftp.sgi.com/graphics/SGIIMAGESPEC">ftp://ftp.sgi.com/graphics/SGIIMAGESPEC</a>

Output resolutions:

Width	Height	Scanning	Frames per second	Comments
3840	2160	progressive	50	Master
1920	1080	progressive	50	
1280	720	progressive	50	
1920	1080	interlaced <sup>5</sup>	25	Top field first
720	576	interlaced <sup>5</sup>	25	Top field first, 16:9 ‘anamorphic’

## 4. List of distributed 10 sec Sequences

Name	Coding difficulty	By SVT assumed quality difference for secondary distribution giving ‘good reception condition’ (ITU-R BT.1122-1)
<b><i>CrowdRun</i></b>	Difficult	<30 %
<b><i>ParkJoy</i></b>	Difficult	<30 %
<b><i>DucksTakeOff</i></b>	Difficult	<30 %
<b><i>IntoTree</i></b>	Easy	<12 %
<b><i>OldTownCross</i></b>	Easy	<12 %

The whole demanding, but not unduly so, multi-genre TV-program is called ‘Fairytale’ (by SVT).

The above sequences are all 10 seconds long.

<sup>1</sup> Details to the scanning process can be found in section 1.1

<sup>2</sup> Details for the process of converting the 10 bit logarithmic data to 16 bit linear can be found in 5.3

<sup>3</sup> Parameters from the ITU-R BT.709 can be found in 5.1

<sup>4</sup> Details for the process of colour correction can be found in 5.4

<sup>5</sup> Details to the process of generating the interlaced formats can be found in 5.5

Progressive sequences have a duration of 500 frames – interlaced sequences have a duration of 250 frames.

The above list is ordered according to the coding difficulty with the most difficult sequence at the top of the list.

If the whole program is encoded at a bit rate where a sequence classified as ‘*difficult*’ does not have a bigger quality drop than 30 % in a DSCQS test, it can be assumed that ‘good reception condition’ (a.k.a. ‘Broadcast Quality’) according to ITU-R BT.1122-1 can be obtained for the whole program.

The same appears for sequences classified as ‘*medium*’ (between 30 % and 12 % quality drop allowed) and sequences classified as ‘*easy*’ (12% quality drop allowed).

The sequences have been designed to represent a demanding, but not unduly so, multi-genre TV program, hence it can be suggested that this set of sequences can act as a reference for the peak bit rate (video elementary stream) that is needed for secondary distribution (a.k.a. transmission) of most TV programs to reach ‘good reception condition’ (a.k.a. broadcast quality).

## 5. Appendix

### 5.1. The ITU-R BT.709 Transfer and Chromaticity parameters

CIE coordinates for ITU-R BT.709 reference primaries and CIE standard illuminant at White Point 6500 Kelvin (a.k.a. D65):

	Red	Green	Blue	White (D65)
x	0.6400	0.300	0.1500	0.3127
y	0.3300	0.600	0.0600	0.3290
z	0.0300	0.100	0.7900	0.3538

Transfer parameters (a.k.a. gamma correction) based on ITU-R BT.709:

$$R, G, B < 0.018$$

$$R' = 4.5 \cdot R$$

$$G' = 4.5 \cdot G$$

$$B' = 4.5 \cdot B$$

$$R, G, B \geq 0.018$$

$$R' = 1.099R^{0.45} - 0.099$$

$$G' = 1.099G^{0.45} - 0.099$$

$$B' = 1.099B^{0.45} - 0.099$$

*Note: the R,G,B values are in the full interval of 0 to (2<sup>16</sup> - 1) inclusive – i.e. no head room or foot room in the supplied ‘sgi16’ data format. Hence the dynamic range has to be ‘squeezed’/‘reduced proportionally’ to legal gamut values when performing conversions to different standardised video formats that uses head room and foot room!*

## 5.2. Details to the film scanning process

The film was scanned using a Filmlight 65mm NorthLight film scanner.

The scanner was set up to scan an area of 50.94mm times 22.34mm of the film. A cropped region was then used to prepare for the post production mastering in 3840x2160 pixels at 50 fps.

The output of the scanner were 'cineon' files with a bit depth of 10 bit (logarithmic scale) for each colour plane.

The Kodak Cineon colour space was used and the following gains were used for the three colour planes:

Color Plane	Gain
Red	1.03
Green	0.96
Blue	0.98

*Note: no interpolation was done during the scanning of the film – each pixel directly corresponds 8.29159 microns of the film.*

## 5.3. Log to linear bit depth conversion

The scanned data was stored using a 10 bit logarithmic scale. This was then converted in the Apple Shake software to full interval 16 bit linear for further processing. For this Log-to-Lin process the highest and the lowest pixel value for each colour plane in every single scene was detected and the value '0' was applied to the lowest value whereas the value '65535' was applied to the highest value to make use of the full interval 16 bit scale for every single scene.

## 5.4. Colour correction and post processing

Colour correction was done in Shake using a 1920x1080p native resolution monitor at viewing distance 3H (three times the monitor height) and calibrated based on ITU-R BT.709. Other artistic post processing steps (such as grain reduction and added perceived 'sharpness') were done by viewing the monitor, in those specific steps, displaying only a cut out 1920x1080 pixel region (1/4 of the full frame with direct pixel to pixel mapping) – still in a distance of 3H, hence assuming a viewing distance of 1.5H for '2160p/50 Super-HD' i.e. the 4xHDTV resolution.

## 5.5. Interlacing

To aid the interlacing process a basic raster of 2164 lines (3840x2164p/50) was used as input.

### a) Interlacing to 1080i

For interlacing, every second 2164p-frame was shifted vertically two lines downwards. After deleting the first two and last two lines in the frames that were shifted (and the last four lines for the frames that were not shifted) to get 2160 lines again, each frame was filtered to 540 lines by line averaging (using Shake's Box Filter). Horizontal filtering from 3840 to 1920 columns was performed using Shake's Sinc Filter to benefit in perceived sharpness from the "oversampled" master.

The two 540-line fields were then weaved into one single 1080-line interlaced frame.

This process resembles the process in any video camera performing interlace in the basic default 'Field Integration Mode' – i.e. like a 2160 line video camera sensor reading out the average of the sensor's line 1+2+3+4 to Field 1; line 3+4+5+6 to Field 2; line 5+6+7+8 to Field 1 etc.

### b) Interlacing to 576i

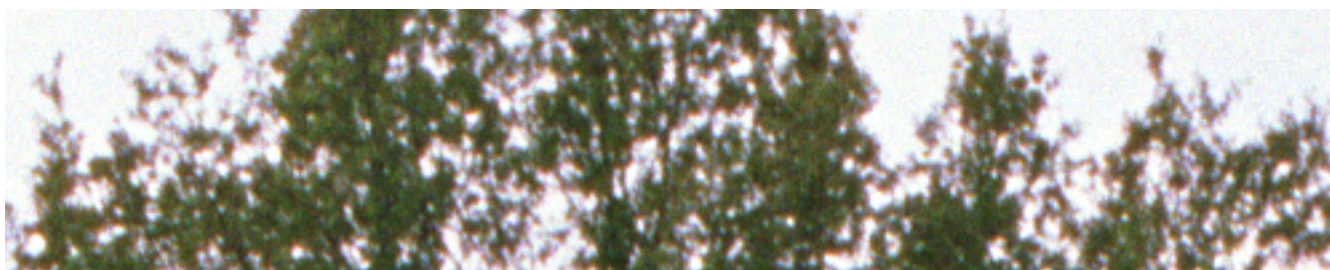
A similar process was applied for interlacing to Standard Definition (576i/25 in 16:9 'anamorphic'). Here the vertical shifting was for 3.75 lines. Line/column averaging (using Sinc Filter horizontally and Box Filter vertically) was then done to reach 288 lines and 702 columns (702 pixels at 13.5 MHz corresponds to 52 us as in PAL, hence correct pixel aspect ratio by legacy). Adding 9 columns of mid-grey at the left and right border a horizontal resolution of 720 columns was reached.

## 5.6. Resolution Comparison

The following series of three images shows a cut out of the top of the first frame of the *CrowdRun* sequence (frame number 7111 in 50p program versions).

Beside the full resolution master an upsampled version of the 720p frame and an upsampled version of the 1080p frame is shown to demonstrate the amount of detail present in the different resolutions. Upsampling was done using a Lanczos filter.

The images clearly show that the 3840x2160 Master has a higher spatial fidelity than what can be reached by any native 1080p/50 video camera.



Cut out from the 3840x2160 Master



Cut out from the downsampled 1920x1080 version (upsampled using a Lanczos filter)



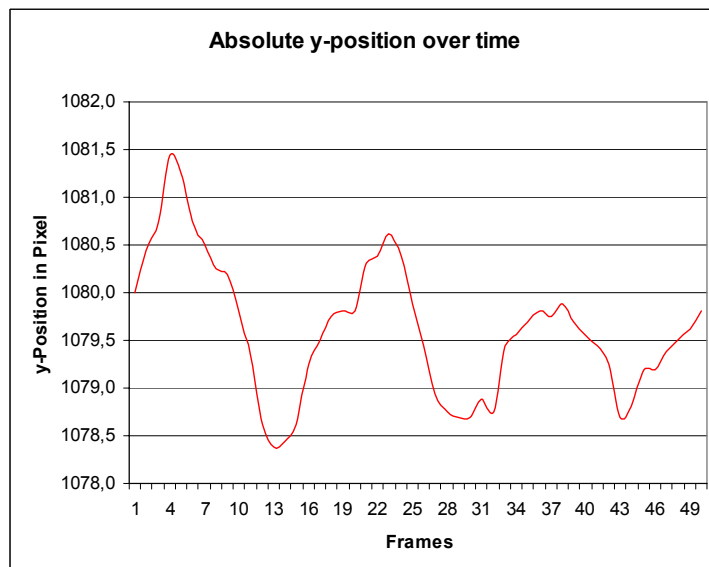
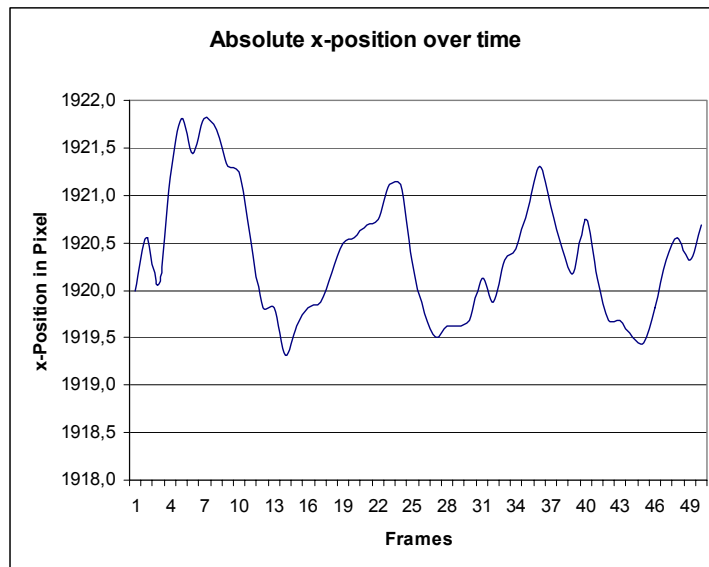
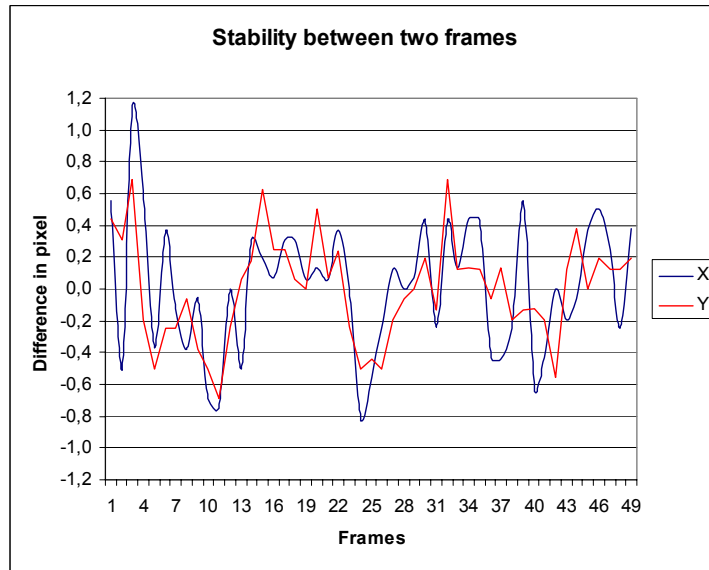
Cut out from the downsampled 1280x720 version (upsampled using a Lanczos filter)

## 5.7. Spatio-temporal stability of the camera system

To determine the temporal stability of the whole capturing system (film camera and scanner) one fixed point on a fixed test-chart with a fixed camera position, in 3840x2160p/50, was tracked for the duration of one second. In a perfect stable system (like in video camera systems) this point would be found at the same position of the scanned image for each single



frame. The graphs below show the stability of the whole system for the 3840x2160p/50 master - downsampled versions have better stability.



## 5.8. Distribution of the whole program

Upon special request and only if one of the contact persons given in 5.10 decide they have resources to help, the whole six minute program 'Fairytale' (by SVT) including 5.1 and 2.0 (48 kHz, 24 bit linear) audio can be made available as disk copy.

As distribution of this huge amount of data requires quite some effort, anybody who received a copy of the whole program is especially asked to make this data available to other parties to reach a super distribution effect.

## 5.9. List of proposed 10 seconds cut outs

Name	Coding difficulty	By SVT assumed quality difference for secondary distribution giving 'good reception condition' (ITU-R BT.1122-1)	First Frame No in whole multi-genre program. 50p-formats totally No 1-18678.	-"- 25i-formats totally No 1-9339.
<b>CrowdRun</b>	Difficult	<30 %	7111	3556
<b>ParkJoy</b>	Difficult	<30 %	15523	7762
PrincessRun	Difficult	<30 %	10429	5215
<b>DucksTakeOff</b>	Difficult	<30 %	12987	6494
Seeking	Medium	<30 % >12 %	11911	5956
TreeTilt	Medium	<30 % >12 %	8825	4413
PassingBy	Easy	<12 %	14131	7066
<b>IntoTree</b>	Easy	<12 %	5119	2560
Umbrella	Easy	<12 %	12463	6232
OldTownPan	Easy	<12 %	427	214
<b>OldTownCross</b>	Easy	<12 %	1217	609
IntoCastle	Easy	<12 %	3777	1889
DanceKiss	Easy	<12 %	17953	8977

*Note: sequences written in bold are the sequences distributed independently from the whole program (see section 1).*

## 5.10. Data Distribution Contacts

Kjell Brunnstrom, Acreo, Sweden – [kjell.brunnstrom@acreo.se](mailto:kjell.brunnstrom@acreo.se)

Hans Hoffmann, The EBU, Geneva (see 5.11) – [hoffmann@ebu.ch](mailto:hoffmann@ebu.ch)

Tobias Oelbaum, Munich University of Technology, Germany – [oelbaum@tum.de](mailto:oelbaum@tum.de)

Lars Haglund, SVT, Sweden – [lars.haglund@svt.se](mailto:lars.haglund@svt.se)

## 5.11. Data Distribution of 4:2:2 version

In addition to the sgi16-data, the EBU is providing 10 bit, 4:2:2 versions as 'yuv16-data' of the sequences, that comply to the RGB to YCbCr encoding parameters and dynamic range legal gamut values in ITU-R BT. 709 (the same as in SMPTE 274M for 1080 line video formats and SMPTE 296M for the 720 line format) and that comply to ITU-R BT. 601 for the 576 line video format. The conversion from sgi16-data to 10 bit 4:2:2-video (as yuv16-data with six padding LSB's) was performed at the EBU with the DVS Clipster Software 2.0.0.52.

<a href="http://www.ebu.ch/en/technical/hdtv/test_sequences.php">http://www.ebu.ch/en/technical/hdtv/test_sequences.php</a>	EBU order form (for EBU members only)
<a href="ftp://vqeg.its.bldrdoc.gov">ftp://vqeg.its.bldrdoc.gov</a>	Free download – ftp server managed by Video Quality Experts group (VQEG), <a href="http://www.its.bldrdoc.gov/vqeg/">http://www.its.bldrdoc.gov/vqeg/</a>

### 5.12. Recommended Viewer

For viewing the original ‘sgi16’ files we recommend the use of XnView. This tool is freeware for non-commercial use (a licence for commercial use can be obtained for a small fee) and can be downloaded at <http://www.xnview.com>. XnView also can be used to convert the original images into other image formats such as tif.

### 5.13. Conversion Software

The command-line tool “sgi2yuv” can be used to convert the ‘sgi16’ data format into various useful raw video bit streams among which are:

- Planar YUV 4:4:4 video data, one file per sequence – extension .yuv
- Planar YUV 4:2:2 video data, one file per sequence – extension .yuv
- ‘Abekas’ YUV 4:2:2 video data, one file per frame – extension .qnt
- Planar YUV 4:2:0 video data, one file per sequence – extension .yuv

The software is currently under development and is available at: <http://www.ldv.ei.tum.de>  
For questions concerning this program please contact: Tobias Oelbaum, [oelbaum@tum.de](mailto:oelbaum@tum.de)

### 5.14. History of this document

Version 1.0 February 2006