

Een nieuwe standaard voor ondertitels bij video

Er is behoefte om de layout van ondertitels gedetailleerd op te kunnen geven in een script. Het meest gebruikte script formaat subrip (extensie .srt) voorziet hierin niet. Vrije uitbreidingen binnen subrip script worden doorgaans niet geaccepteerd / herkend door programma's die subrip verwerken.

Er bestaat een ondertitel formaat waarmee enkele tientallen layout opdrachten kunnen opgegeven, het heeft de naam substation alpha script (extensie .ass). Het bestandsformaat voor de mens moeilijk leesbaar en kent technische onduidelijkheden en beperkingen.

Om die reden heb ik de afgelopen week gewerkt aan een nieuwe standaard die uitgaat van het populaire subrip, die toegang biedt tot alle layout mogelijkheden van substation alpha. Dit zijn de overwegingen:

Specificatie van titel en commentaar

In de praktijk blijkt dat ondertitel scripts soms niet bij de juiste film / documentaire worden bewaard. Vandaar dat in het nieuwe script formaat een titel specificatie is opgenomen. Soms is behoefte van maker of gebruiker om commentaar toe te voegen, dat kan in het nieuwe formaat.

Definitie van layout stijlen

Analoog met de cascade style sheet specificaties voor html code, kent het nieuwe script een of meerdere stijlen. Minimaal is er een default stijl. Alle stijlen worden gespecificeerd aan het begin van het script. De specificatie is kort en bondig, elke stijl specificatie past op één regel. Waar in het ass-formaat een stijl is gedefinieerd als een reeks van getallen, zijn de SrtPlus stijlen zoveel mogelijk herkenbaar verwoord.

Beeldmerk

Volgend op de stijl specificatie volgen de tekst-items. Een tekst item begint met twee tijdcodes, elk in het formaat "`\d\d:\d\d:\d\d:\d\d`", de digit velden zijn voor de uren, minuten, seconden en de honderdsten van seconden. Tussen de tijdcodes is voor de leesbaarheid hetzelfde beeldmerk als voor subrip, de tekenreeks " -> ". Op dezelfde regel als de tijdcodes volgt optioneel de naam van de stijl die op het item van toepassing is. Indien geen stijl is opgegeven, is de default stijl van toepassing. Op deze regel volgen één of meer tekstregels in het unicode formaat. De regels worden gescheiden door de bekende carriage return en/of linefeed tekens. Op elk item volgt een blanco regel, ten behoeve van de leesbaarheid.

Centi-seconden in plaats van milliseconden

In subrip zijn tijdstempels met milliseconden. Het menselijk waarnemingsvermogen voor geluidstiming houdt op bij tienden van seconden, in de nieuwe standaard zijn slechts twee digits achter de comma opgenomen.

Geen item-nummers

Subrip kent een item-nummering voor referentie bijvoorbeeld in het geval van syntax-fouten. In plaats daarvan refereren aan script-regelnummers is softwarematig bij te houden. Bovendien, voor diegenen die scripts handmatig bewerken, bestaan editor programma's die regelnummers tonen. Dank aan Lea Bouwmeester voor het advies.

Jump start

Een nieuw formaat ontwerpen zou slechts luchtftietserij zijn indien het ontworpen script niet direct bruikbaar is voor video afspeler software voor computer / mobiel en voor moderne televisies. Ten behoeve van een snelle introductie is van belang dat het nieuwe script geconverteerd kan naar een scriptformaat dat reeds nu wordt ondersteund. Om die reden is deze week (windows) software gepubliceerd op Github.

Conversie software

Enkele kwaliteitskenmerken van het SrtPlus conversie programma:

- A. Het herkent scripts in het srt, ass en srtplus formaat. Voor de uitvoer kan gekozen worden uit de zelfde drie formaten.
- B. Kleurcodes van ass files zijn in het legacy &HAARRGGBB formaat en worden geconverteerd naar de van html bekende kleurcodes zoals #RGB #RGBA #RRGGBB en #RRGGBBAA. De conversie is vice versa, dus van SrtPlus naar ass is weer een terug codering.
- C. Alle gegevens elementen van ass en srt worden zonder verlies overgenomen naar srtplus. Omgekeerd gaan geen gegevens verloren bij terug coderen naar ass.
- D. Bij de conversie van ass stijlnamen gaan de ass stijlnamen voor. Er is dus nooit een conflict met de ingebouwde SrtPlus stijlen.
- E. Bij import is een toets op de validiteit van de scripts, teneinde het voorkomen van fouten bij het afspelen.

Afgelopen week is nog niet gelukt:

1. Subrip kent enkele inline (binnen de tekstregels) coderingen voor cursief en dergelijken. Deze worden in de huidige versie van SrtPlus.exe onverkort overgenomen. De bedoeling zoveel mogelijk de layout per item te regelen, middels een stijl. Daar waar meerdere stijlen binnen een item nodig zijn, is de suggestie van Dominique van der Heyden: het adapteren van de goed gedefinieerde ass tags binnen de SrtPlus standaard.
2. Een sterke test op de validiteit van invoer scripts.
3. Het SrtPlus gebeuren is een onderdeel van het CinemaVision project en is nog niet binnen CinemaVision geïntegreerd.

Gecorrespondeerd is afgelopen week met Wouter Bos van NL-Invest voor resources ten behoeve van verdere ontwikkeling.

Discussie

SrtPlus kent nu twee tijdstempels per item, de begin- en eindtijd. Het is mogelijk uit te komen met alleen de starttijd en uit te gaan van:

- Een minimale leessnelheid van standaard 12 tekens per seconde.
- De tijd van vertonen van een item is in snelle dialogen begrensd door de start van het volgende item.

Op basis van deze regels kan een computer de eindtijd berekenen. Het weglaten van de eindtijd uit SrtPlus files vereenvoudigt de codering. Dank aan Martin van Rijn.

Graag uw opinie over de nieuwe standaard.

Bert Kerkhof
kerkhof.bert@gmail.com

Bijlagen: een voorbeeld ondertiteling in drie verschillende script formaten

Bijlage A: voorbeeld ondertitel in Subrip script:

1

00:01:49,400 --> 00:01:53,230

943-66-2291,
Afdeling Pensioenen?

2

00:01:54,860 --> 00:01:58,280

<I>Mr Moses, uw contactpersoon<\I>
<I>komt eraan.<\I>

3

00:02:15,340 --> 00:02:18,330

Hallo, met Sarah.
- Met Frank Moses.

4

00:02:18,930 --> 00:02:21,250

Frank, hoe gaat het?

5

00:02:21,970 --> 00:02:23,300

Stil.

6

00:02:25,140 --> 00:02:28,960

Ik bel omdat ze die cheque
weer zijn vergeten.

7

00:02:30,230 --> 00:02:33,310

Ongelofelijk dat ze 't nog
steeds niet hebben geregeld.

8

00:02:33,310 --> 00:02:35,480

Ik stuur een nieuwe.

9

00:02:35,690 --> 00:02:39,680

<I>Het spijt me vreselijk.<\I>
- Ach, wat doe je eraan?

10

00:02:41,610 --> 00:02:43,860

Hoe gaat 't met je avocado?

Bijlage B: voorbeeld ondertitel in SrtPlus script:

[SrtPlus v1 styles]

; (c) SrtPlus is designed by Bert Kerkhof

Title: Red (2010).dutch

Default = Font Arial Size 20 White Transparent Outline 2 Shadow 1 Bottom Center

!SH = Layer 1

[Events]

00:01:49,40 --> 00:01:53,23

943-66-2291,

Afdeling Pensioenen?

00:01:54,86 --> 00:01:58,28 Italic

Mr Moses, uw contactpersoon

komt eraan.

00:02:15,34 --> 00:02:18,33

Hallo, met Sarah.

- Met Frank Moses.

00:02:18,93 --> 00:02:21,25

Frank, hoe gaat het?

00:02:21,97 --> 00:02:23,30

Stil.

00:02:25,14 --> 00:02:28,96

Ik bel omdat ze die cheque

weer zijn vergeten.

00:02:30,23 --> 00:02:33,31

Ongelofelijk dat ze 't nog

steeds niet hebben geregeld.

00:02:33,31 --> 00:02:35,48

Ik stuur een nieuwe.

00:02:35,69 --> 00:02:39,68

{\i1}Het spijt me vreselijk.{\i0}

- Ach, wat doe je eraan?

00:02:40,00 --> 00:02:41,00 !SH

Zucht...

00:02:41,61 --> 00:02:43,86

Hoe gaat 't met je avocado?

[End of script]

Bijlage C: voorbeeld ondertitel in ass script:

[Script Info]

; This is an Advanced Sub Station Alpha v4+ script.

Title: Red (2010).dutch

ScriptType: v4.00+

Collisions: Normal

PlayDepth: 0

[V4+ Styles]

Format: Name, Fontname, Fontsize, PrimaryColour, SecondaryColour, OutlineColour, BackColour, Bold, Italic, Underline, StrikeOut, ScaleX, ScaleY, Spacing, Angle, BorderStyle, Outline, Shadow, Alignment, MarginL, MarginR, MarginV, Encoding

Style: Default,Arial,20,&H00FFFFFF,&H0300FFFF,&H00000000,&H02000000,0,0,0,100,100,0,0,1,2,1,2,0,0,0,1

[Events]

Format: Layer, Start, End, Style, Actor, MarginL, MarginR, MarginV, Effect, Text

Dialogue: 0,0:01:49.40,0:01:53.23,Default,,0,0,0,,943-66-2291,\NAfdeling Pensioenen?

Dialogue: 0,0:01:54.86,0:01:58.28,Default,,0,0,0,,{\i1}Mr Moses, uw contactpersoon{\i0}\N{\i1}komt eraan.{\i0}

Dialogue: 0,0:02:15.34,0:02:18.33,Default,,0,0,0,,Hallo, met Sarah.\N- Met Frank Moses.

Dialogue: 0,0:02:18.93,0:02:21.25,Default,,0,0,0,,Frank, hoe gaat het?

Dialogue: 0,0:02:21.97,0:02:23.30,Default,,0,0,0,,Stil.

Dialogue: 0,0:02:25.14,0:02:28.96,Default,,0,0,0,,Ik bel omdat ze die cheque\Nweer zijn vergeten.

Dialogue: 0,0:02:30.23,0:02:33.31,Default,,0,0,0,,Ongelofelijk dat ze 't nog\Nsteeds niet hebben geregeld.

Dialogue: 0,0:02:33.31,0:02:35.48,Default,,0,0,0,,Ik stuur een nieuwe.

Dialogue: 0,0:02:35.69,0:02:39.68,Default,,0,0,0,,{\i1}Het spijt me vreselijk.{\i0}\N- Ach, wat doe je eraan?

Dialogue: 0,0:02:41.61,0:02:43.86,Default,,0,0,0,,Hoe gaat 't met je avocado?