PHILIPS

Philips Interactive Media

Technical Note #104

Aspect Ratios in CD-i

Andrew Duncan
Advanced Development
Philips Interactive Media

July 21, 1994

Abstract

There is a great deal of misunderstanding about aspect ratios in creating graphics and video for use in CD-i applications. This note attempts to quantify these issues to provide guidelines for preparation of the visual assets used in a CD-i title.

Publication History:
No revisions.

Note that TN numbers 91 through 100 are reserved for conversion
of Philips TSA numbers 1 through 10 to the technical note catalog.

To receive Philips Interactive Media technical notes and other publications, or for
more information, please contact the person below who is designated for your area.

From Europe and the Middle East:

Hein Zegers
Developer Support
Philips Interactive Media Centre
Maastrichterstraat 63
B-3500 Hasselt
Belgium
Fax: +32 11 242168
Internet ID: hein@pimc.be

From the USA and Asia:

Lucy Lediaev
Developer Support
Philips Interactive Media
11050 Santa Monica Boulevard
Los Angeles, CA 90025
USA
Fax: +1 310 477 4953
Internet ID: lucy@aimla.com
CompuServe ID: 72056,1130

INTRODUCTION

The whole question of aspect ratios for CD-i is filled with misunderstandings, many of them "just" semantic, but of course, whatever the cause, the result is confusion. First, some definitions.

quantity	definition	symbol
physical dimensions	Height and width of screen image in cm.	$Screen_{x,y}$
logical dimensions	Height and width of screen image	$Pixels_{x,y}$
pixel size: $Size_{x,y}$	Measured in cm/pixel.	$Pixel$

Clearly,
$$Screen_x = Pixel\ Size_x \ [\ Pixels_x \quad \text{(and likewise in the y-direction). [Eq. 1]}$$
$$(cm) = (cm/pix) \ [\ (pix)$$

Taking the ratio of y to x:

$$\frac{Screen_y}{Screen_x} = \frac{Pixel\ Size_y}{Pixel\ Size_x} \ [\ \frac{Pixels_y}{Pixels_x} \qquad \text{[Eq. 2]}$$

physical (display) ratio (often 0.75)	pixel (aspect) ratio	logical (data) ratio

All three of these ratios are called *the* aspect ratio in various contexts! Let's keep them separate. Hereafter we will drop the word "aspect" from the term "pixel ratio".

> *Note: One may model pixels as points or rectangles. If they are points, then what we call "pixel size" should properly be called "pixel spacing." In addition, Eq. 1 should read Screen = Pixel Spacing [(Pixels - 1), because the image does not extend to the left of the first pixel or to the right of the last. (See Fig. 1)*

```
left- 1          right- 4                      left- 1          right- 4

top- 1                                   top - 1   ●   ●   ●   ●   ┬
                                                                  height- 1
bottom- 2                                bottom- 2  ●   ●   ●   ●   ┴

          width- 4                               ├─ width- 3 ─┤

   Modeling  pixels  as  rectangles          Modeling  pixels  as  points
height- # of pixels vertically - bottom- top+ 1   height- # of pixels vertically - 1 - bottom- top
width- # of pixels horizontally - right-left + 1  width- # of pixels horizontally - 1 - right- left
```
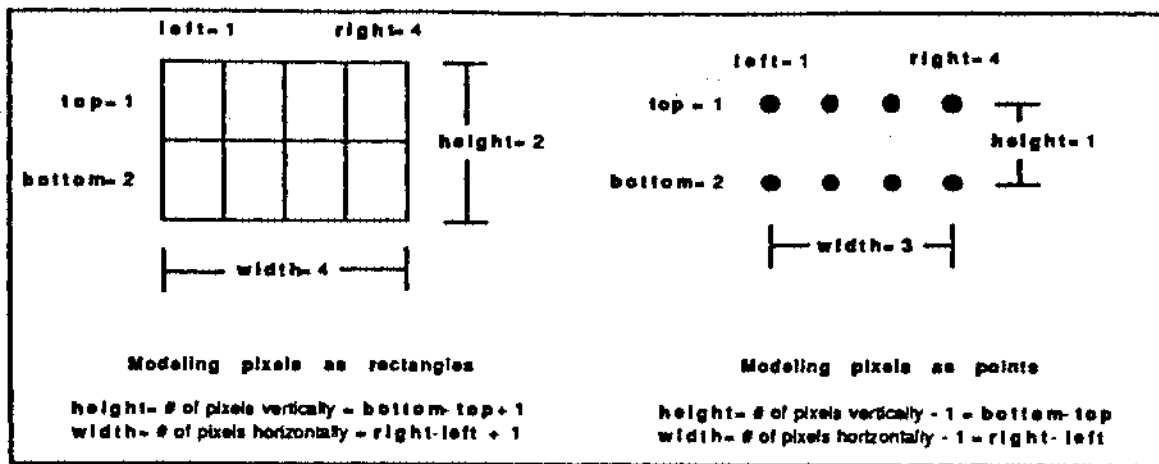
Figure 1

In this document, we will think of pixels as rectangles with non-zero height and width. The alteration needed in subsequent equations is trivial, and, in fact, the difference it makes becomes negligible for large images.

IMAGE CONVERSION

Differing pixel ratios have to be accounted for when transferring artwork from one medium to another. First consider what will happen if an image is generated on a computer and then transferred pixel-for-pixel directly to CD-i. Computer displays typically have a pixel ratio close to 1. The NTSC CD-i pixels are taller, with a ratio of 1.23 to one computer pixel. Figure 2 shows what happens if no image correction is applied: what was a circle to the artist at the computer screen is a tall ellipse to the CD-i viewer.

How do we arrive at a good-looking CD-i display? Before we develop a general equation, we can work backward from the desired result to the starting image that will yield that result. In Figure 3, we have a CD-i display with a circle that looks properly round. (Artwork limitations make it look a bit tall. Consider it round.) Moving back to the computer image that generates it, we see that the "circle" must look short. In fact, everything in the artwork must look vertically squashed. Well, this is not a convenient way for artists to work. How can the artist work in a domain where things look to his or her eye the way they will appear on the final CD-i display? This is where resampling comes in.
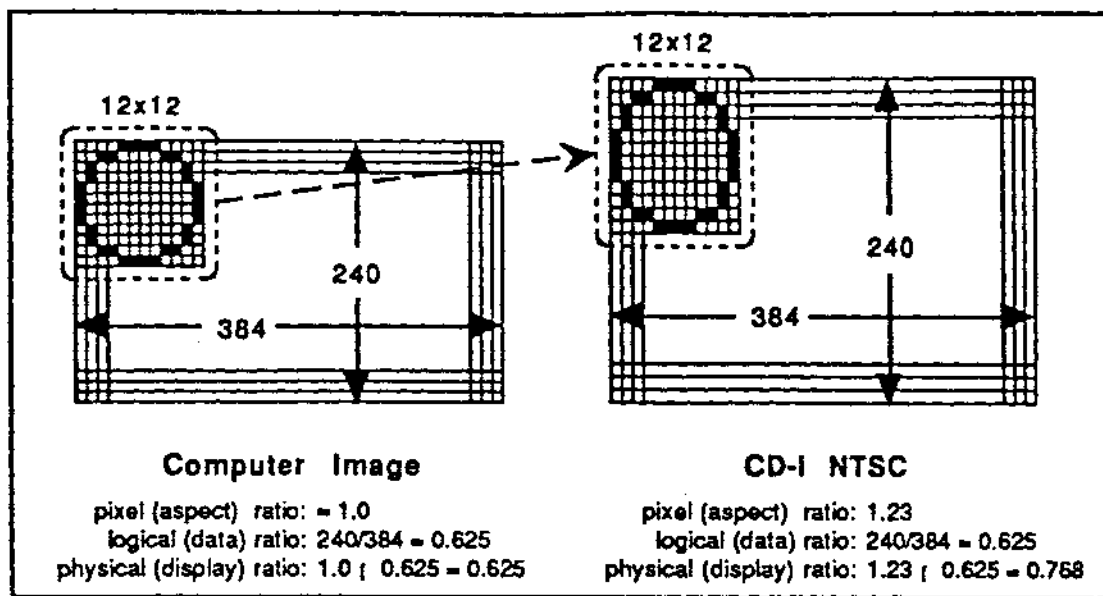
Computer Image

pixel (aspect) ratio: = 1.0
logical (data) ratio: 240/384 = 0.625
physical (display) ratio: 1.0 ¡ 0.625 = 0.625

CD-I NTSC

pixel (aspect) ratio: 1.23
logical (data) ratio: 240/384 = 0.625
physical (display) ratio: 1.23 ¡ 0.625 = 0.768
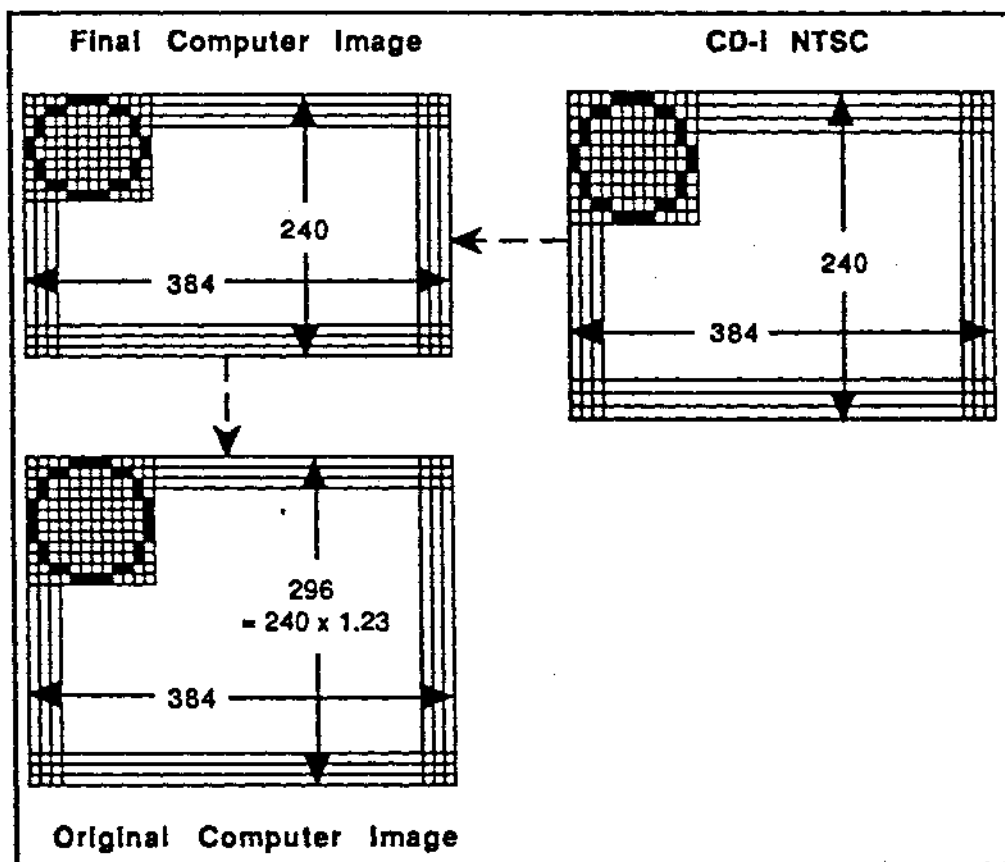
Figure 2

Figure 3

The original artwork can be created in a taller size and then squeezed to the proper size. In the process of squeezing, the image will be pre-distorted in so that it will look right again when displayed on the CD-i TV screen. The logical height of the original image should be the height of the final image multiplied by the pixel ratio of the final image. (The alert reader will note that the product has been rounded up to an even number, out of sheer obsessive compulsion.) The general equation that gives the required starting height is Eq. 6b below.

How is this squeezing accomplished? Resampling can be described as the process of reconstructing a smooth curve that describes the variation of brightness and color from pixel to pixel and then finding values on that curve in between the original ones. This is illustrated in the figure below.
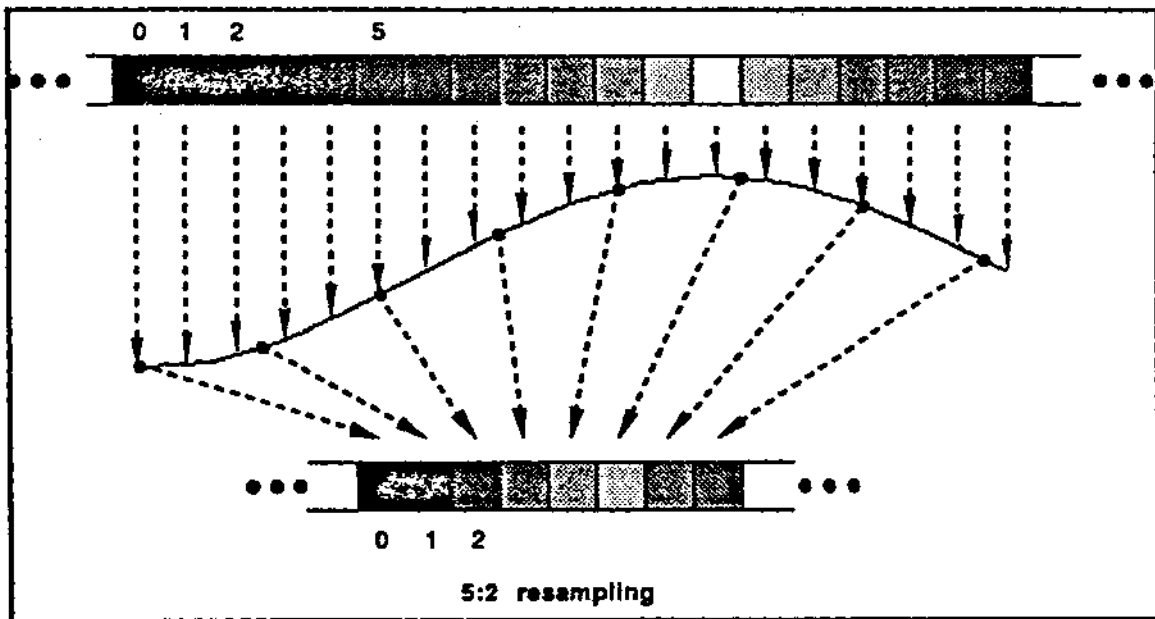

5:2 resampling

Figure 4

It is important to recognize that resampling and pixel ratio change are identical. One cannot resample an image (for instance, to convert from a logical size of 200x100 to a logical size of 100x200) and *then* do aspect ratio correction to make circles look round again. Once you have resampled, the *only* remaining factor that bears on the proportion of the final image displayed is the pixel ratio of the play-back device.

For example, imagine someone asks us to convert a 6x4 image into a 12x4 image. We may do this by sampling the source image at twice the resolution horizontally, as shown in Fig. 5. We obtain the desired 12-horizontal-pixel format, but when the image is displayed, the colored square has become elongated. Resampling has caused, and is identical to, pixel ratio distortion. To fix this, we would have to change the pixel ratio of the destination display, grab from a shorter rectangle in the source image, or display to a narrower rectangle in the destination display. In the real world, the first option is not available. The second option involves cropping some of the original image— which may not be acceptable, for example in a feature film. The third option does not use all the available display area; this is known as "letterboxing."
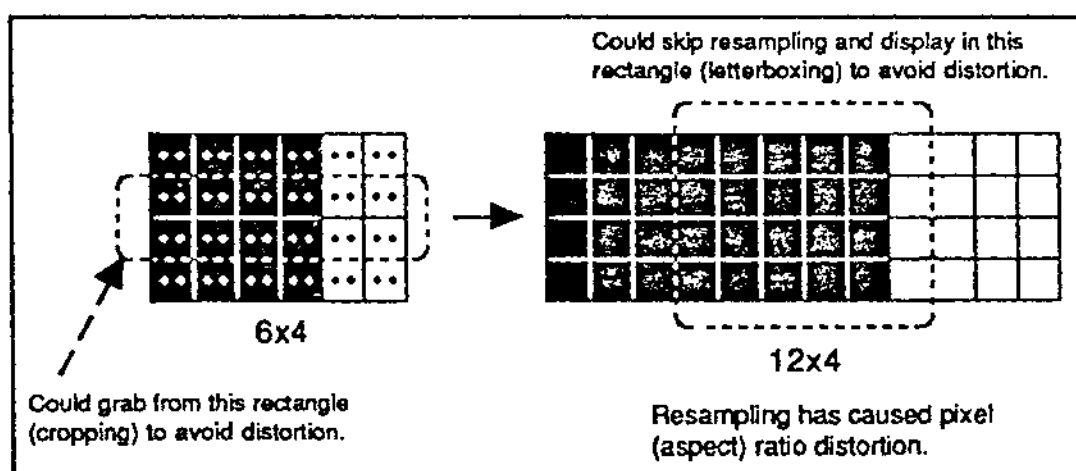

Figure 5

THE GENERAL EQUATION

Without clearly defined terminology, it is almost impossible to discuss the issues raised here. Defining the terms:

QUANTITY	DEFINITION	SYMBOL
source pixel ratio	Ratio of pixel size (or spacing) dy/dx	R_s
source logical grab dimensions	Height & width in pixels of region to be resampled	$Grab_{x,y}$
result logical dimensions	Height & width in pixels of resampled image	$Result_{x,y}$
playback pixel ratio	Pixel ratio on playback device	R_p
pixel ratio distortion	> 1 if playback is too skinny; < 1 if too fat.	D

Consider what the *physical* display dimensions of the grab rectangle are. From Eq. 2 the ratio of its height to width must be

$$\text{grab rect physical ratio} \quad = \quad R_s \mid \frac{Grab_y}{Grab_x} \, . \qquad \text{[Eq. 6a]}$$

This rectangle may have to be squeezed or stretched (resampled) to make it fit into the playback rectangle, whose proportion is

$$\text{playback rect physical ratio} \quad = \quad R_p \mid \frac{Result_y}{Result_x} \, . \qquad \text{[Eq. 6b]}$$

It is the *ratio* of these two expressions that describes the distortion of the image. If they are equal, a figure that looked square when the original material was played back on the source hardware will still look square on the playback hardware; circles will not turn into ovals, etc. Thus to convert an image from a source pixel ratio of R_s to display properly on a device with pixel ratio R_p, we must have

$$R_s \mid \frac{Grab_y}{Grab_x} = R_p \mid \frac{Result_y}{Result_x} \qquad \text{[Eq. 7a]}$$

which is the same as

$$Grab_y = \frac{R_p \mid Result_y}{R_s \mid Result_x} \mid Grab_x \qquad \text{[Eq. 7b]}$$

which is the same as

$$R_p = R_s \mid \frac{Grab_y/Grab_x}{Result_y/Result_x} \, . \qquad \text{[Eq. 7c]}$$

Notice that Eq. 7b gives the starting (grab) height of an image to be converted, as illustrated in Fig. 3 and described above. Eq. 7c is often described by saying that the pixel ratio is being *converted* from R_s to R_p. The explicit solution for the distortion D (just the ratio of Eq. 6a and 6b) is of less immediate value. Consider, though, that the expression for D is a ratio of ratios of ratios. Discussions of the value or meaning of D usually flounder in misunderstanding.

PIXEL RATIOS FOR CD-i BASE CASE

CD-i pixel ratios are determined by the video hardware in the CD-i player. The Green Book says (§V.2.4.1):

$$pixel\ ratio = \frac{pixel\ height}{pixel\ width} = \frac{pixel\ clock\ frequency\ [\ scan\ line\ period}{\#\ of\ active\ lines} \qquad [\ \frac{3}{4}.$$

[Eq.3]

How do we understand this in terms of the foregoing? What's the 3/4 all about? Observe that

$$Pixels_x = pixel\ clock\ frequency\ [\ scan\ line\ period \qquad [Eq.\ 4]$$

$$Pixels_y = \#\ of\ active\ lines$$

Substituting that (and pretending the approximate equalities are exact) yields

$$pixel\ ratio = \frac{Pixels_x}{Pixels_y} \qquad [\ \frac{3}{4}.$$

or
$$3/4 = pixel\ ratio\ [\ logical\ ratio. \qquad [Eq.\ 5]$$

This is the same as Eq. 2. Thus, this equation says that the CD-i player draws a picture on the screen whose proportions are 3:4 in *all cases*.

To further confuse the issue, the numbers for pixel clock, scan line period, and pixel ratio in §V.2.4 of the Green Book are incorrect. The Green Book says to use 242.5 as the active line count for NTSC and 287.5 for PAL, instead of 240 and 280 respectively. Similarly, the product of pixel clock rate and active scan line period gives a pixel count larger than 384, or its high-res double value of 768.

To further confuse the issue, the numbers for pixel clock and scan line period in §V.2.4 of the Green Book are incorrect. The correct numbers are:

format	pixel clock frequency	scan line period	pixel ratio	
525 TV	15.1049 MHz	52.6555 µs	1.230	[Green Book: 1.19]
625	15.0 MHz	52.0 µs	1.017	[Green Book: 1.05]

We have empirically measured the NTSC ratio to be 1.235 ± 0.003 and the PAL/SECAM ratio to be 1.019 ± 0.003.

PIXEL RATIOS FOR DV

The Green Book DV extension discusses the proper pixel ratios for playback of MPEG video in § IX.4.3.2.3. "Full Motion decoders *do not* compensate for aspect ratio distortion." This means that the decoder does not in fact use the information in the aspect ratio field of the MPEG Video Sequence Header. Thus the following sentence, "It is recommended therefore to select one of the following aspect ratio options: ...", *should* mean:

> Before encoding, correct the (pixel) aspect ratio of the digital image to be one of the following ratios, and set the pel aspect ratio index field of the Sequence header correspondingly.

In fact, this is what the Green Book intends the passage to mean; unfortunately the pixel ratios given in the Green Book are wrong. The recommended ratios are:

> "1.19 [1.2015] if no aspect ratio distortion is to occur in decoders producing a 525-line output; in this case, 625-line decoders will have approximately 13% distortion.
>
> 1.05 [1.0695] if no aspect ratio distortion is to occur in decoders producing a 625-line output; in this case, 525-line decoders will have approximately 12% distortion.
>
> 1.12 [1.0950] if aspect ratio are to be equalized over both types of decoders."

(The MPEG standard does not provide for these exact aspect ratios, so the values in brackets should be used.)

We see that the given ratios are the same as the (incorrect) base case ratios. The correct ratios—the pixel ratios produced by the CD-i DV decoder board—are the same as the correct base case ratios. This is obvious (after the fact) when one realizes that the DV cartridge and the base-case video both use the same pixel clock. Thus, video that is being prepared for DV should be converted to pixel ratios of 1.23 for NTSC, 1.017 for PAL, or the geometric mean, 1.118, for a compromise.

The pixel ratios for Video CD (White Book) are precisely 0.9 times those for CD-i. Hence the values are 1.1069 for NTSC and 0.9157 for PAL/SECAM. These numbers are supposed to be identical to the D1/CCIR-601 pixel ratios (see below); the White Book specifies them as such. The CD-i NTSC value is a slight deviation from the standard.

RESAMPLING FROM D1

We will deal here with resampling an image in D1 (CCIR 601) format for playback on a Philips CD-i player with a DV MPEG decoder. Our main constraint is the amount of display area allowable. An MPEG macroblock is a 16x16 pixel region. At a picture rate of 30 Hz, the maximum picture size is 330 macroblocks, the size of a 352x240 rectangle. At 24 or 25 Hz, the picture may be up to 396 macroblocks, the size of a 352x288 rectangle. Note that these are area maxima, and taller or wider rectangles may be used if their areas conform to the area restriction. For Video CD, the picture sizes are restricted to 352x240 for NTSC and 352x288 for PAL/SECAM.

The format data for D1 are as follow:

format	logical dim.	logical dim.	pixel clock	pixel ratio
NTSC	720x486	712x486	13.5 MHz	1.095
PAL/SECAM	720x576	703x576	13.5 MHz	0.9157

Following is a list of recommended resampling choices for various combinations of D1 source and CD-i playback formats. The horizontal and vertical resampling ratios are given for each example, as well as the resulting pixel ratio. The numbers that correspond to digital filter complexity are the denominators (number following the colon) in the resampling ratio. Picture distortions of less than 2% are typically not noticeable. Distortions of 4% will be noticed by film professionals. Distortions of >6% will be noticed by a lot of people.

NTSC/24 D1 → NTSC CD-i

$$\begin{array}{c} 691 \\ \times \\ 480 \end{array} \quad \begin{array}{c} 9:5 \\ \longrightarrow \\ 2:1 \end{array} \quad \begin{array}{c} 384 \\ \times \\ 240 \end{array} \qquad R_p: 1.217 \ (1.1\% \ \text{distortion})$$

$$\begin{array}{c} 713 \\ \times \\ 476 \end{array} \quad \begin{array}{c} 13:7 \\ \longrightarrow \\ 17:8 \end{array} \quad \begin{array}{c} 384 \\ \times \\ 224 \end{array} \qquad R_p: 1.253 \ (1.9\% \ \text{distortion})$$

NTSC/29.97 D1 → NTSC CD-i

$$\begin{array}{c} 690 \\ \times \\ 480 \end{array} \quad \begin{array}{c} 15:8 \\ \longrightarrow \\ 15:7 \end{array} \quad \begin{array}{c} 368 \\ \times \\ 224 \end{array} \qquad R_p: 1.251 \ (1.7\% \ \text{distortion})$$

$$\begin{array}{c} 690 \\ \times \\ 480 \end{array} \quad \begin{array}{c} 15:8 \\ \longrightarrow \\ 19:9 \end{array} \quad \begin{array}{c} 368 \\ \times \\ 224 \end{array} \qquad R_p: 1.233 \ (0.2\% \ \text{distortion})$$

PAL/SECAM D1 → PAL/SECAM CD-i

$$
\begin{array}{ccc}
690 & 15:8 & 368 \\
\times & \xrightarrow{\hspace{1cm}} & \times \\
574 & 19:9 & 272
\end{array}
\qquad R_P\text{: } 1.031 \text{ (1.3\% distortion)}
$$

$$
\begin{array}{ccc}
699 & 19:10 & 368 \\
\times & \xrightarrow{\hspace{1cm}} & \times \\
574 & 19:9 & 272
\end{array}
\qquad R_P\text{: } 1.017 \text{ (0\% distortion)}
$$

NTSC D1 → NTSC Video CD

$$
\begin{array}{ccc}
702 & 2:1 & 352 \\
\times & \xrightarrow{\hspace{1cm}} & \times \\
480 & 2:1 & 240
\end{array}
\qquad R_P\text{: } 1.095 \text{ (1.0\% distortion)}
$$

PAL/SECAM D1 → PAL/SECAM Video CD

$$
\begin{array}{ccc}
702 & 2:1 & 352 \\
\times & \xrightarrow{\hspace{1cm}} & \times \\
576 & 2:1 & 288
\end{array}
\qquad R_P\text{: } 0.9157 \text{ (0\% distortion)}
$$

NTSC/29.97 D1 @ 1.85:1 letterbox → NTSC CD-i

$$
\begin{array}{ccc}
713 & 13:7 & 384 \\
\times & \xrightarrow{\hspace{1cm}} & \times \\
371 & 19:9 & 176
\end{array}
\qquad R_P\text{: } 1.244 \text{ (1.2\% distortion)}
$$

PAL/SECAM D1 @ 1.85:1 letterbox → PAL CD-i

$$
\begin{array}{ccc}
713 & 13:7 & 384 \\
\times & \xrightarrow{\hspace{1cm}} & \times \\
436 & 21:10 & 208
\end{array}
\qquad R_P\text{: } 1.035 \text{ (1.8\% distortion)}
$$