# The 4:2:2/4:2:0 Perfect Reconstruction Filter Set and Its Application in HD-SNG

By A.Nakagawa (FUJITSU LABORATORIES LTD.)

## Abstract

For high-definition satellite newsgathering (HD-SNG) applications, 4:2:0 video codecs are widely used. However, there is one major drawback to 4:2:0 chroma sampling. When two or more 4:2:0 codecs are concatenated, the chroma component gradually becomes blurred because the codecs are usually connected via HD-SDI (a 4:2:2 interconnect) resulting in up-sampling and down-sampling of the chroma component. In order to solve this problem, we have designed a new chroma down-sampling and up-sampling filter set for the 4:2:0 codec whereby chroma degradation is completely suppressed after several stages of concatenation. In this presentation, we explain the features of the 4:2:2/4:2:0 perfect reconstruction filter set and discuss its application for HD-SNG.

## Introduction

A basic requirement of today's broadcast environment is to transmit high-quality pictures at low bit rates in order to optimize the bandwidth of high-definition satellite newsgathering (HD-SNG) and terrestrial microwave links and minimize operating costs. To this end, the adoption of MPEG-4 AVC is advancing worldwide. For HD-SNG and microwave link applications, 10 Mb/s or less is widely used from the viewpoint of picture quality versus channel bandwidth.

Regarding MPEG-4 AVC chroma sampling formats, two kinds are used for broadcast applications, 4:2:2 and 4:2:0. There are tradeoffs between these two chroma formats. It is widely accepted that for lower bit rates, 4:2:0 yields a higher picture quality, whereas for higher bit rates, 4:2:2 is better. 4:2:2 sampling is widely used in archiving applications, sports contribution feeds, and network distribution in which higher channel bandwidths are available. On the other hand, 4:2:0 sampling is used for SNG feeds because of limited and costly satellite bandwidth.

However, there is one major drawback to 4:2:0 chroma sampling. When two or more 4:2:0 codecs are concatenated, the chroma component gradually becomes blurred because the codecs are usually connected via HD-SDI (a 4:2:2 interconnect), resulting in upsampling and downsampling of the chroma component.

To solve this problem, a new chroma downsampling and upsampling filter set has been designed for the 4:2:0 codec, whereby chroma degradation is completely suppressed even after several stages of concatenation. The filter set has the following features:

At first, the chroma downsampling and upsampling filter set satisfies the special condition of "Perfect Reconstruction Filter." This is a key feature for non-degraded 4:2:0.

Moreover, the new 4:2:2/4:2:0 filter set is compatible with conventional 4:2:0 decoders. The phase condition of the new filter set is almost the same as that of a conventional 4:2:0/4:2:2 filter set. That is, chroma displacement will never occur when the 4:2:0 bit stream of the new 4:2:2/4:2:0 filter is decoded by the conventional 4:2:0 decoder. This is a very important feature because it ensures interoperability with third-party decoders that do not implement the new filter set. Furthermore, any encoder or decoder can easily implement the filter set, regardless of its compression technology (H.264, MPEG-2, JPEG 2000 etc.).

The new 4:2:2/4:2:0 filter set on hardware has been implemented. Testing has shown that chroma does not blur even at up to 16 stages of 4:2:0 codec concatenation.

This presentation explains the features of the 4:2:2/4:2:0 perfect reconstruction filter set and its application for HD-SNG are discussed.

## Non-degraded 4:2:0 Features

### 4:2:2 and 4:2:0 Formats

Regarding the chroma sampling formats, two kinds are used for broadcast applications, 4:2:2 and 4:2:0. Figure 1 shows the constitution of 4:2:2 and 4:2:0 formats. An HD picture has 1920 pixel width and 1080 pixel height. Usually, colors are converted to one luminance component called the “Y” component, and two chroma components called “U” and “V.” Because the sensitivity of chroma components in the human visual system is lower than that of luminance, the resolution of chroma components are reduced, compared to the luminance component. The difference of 4:2:2 and 4:2:0 is the degree of chroma resolution reduction.

In the 4:2:2 format, the size of the Y component is the same as the original picture size. Two chroma components are downsampled by a factor of two in the horizontal direction. In the HD format, the size of the Y component is 1920 pixel width and 1080 pixel height, and the size of the chroma components is 960 pixel width and 1080 pixel height.

In the 4:2:0 format, the size of the Y component is also the same as the original picture size. Two chroma components are downsampled by a factor of two, not only in the horizontal direction, but also in the vertical direction. In the HD format, the size of the Y component is also 1920 pixel width and 1080 pixel height, and the size of the chroma components is 960 pixel width and 540 pixel height.

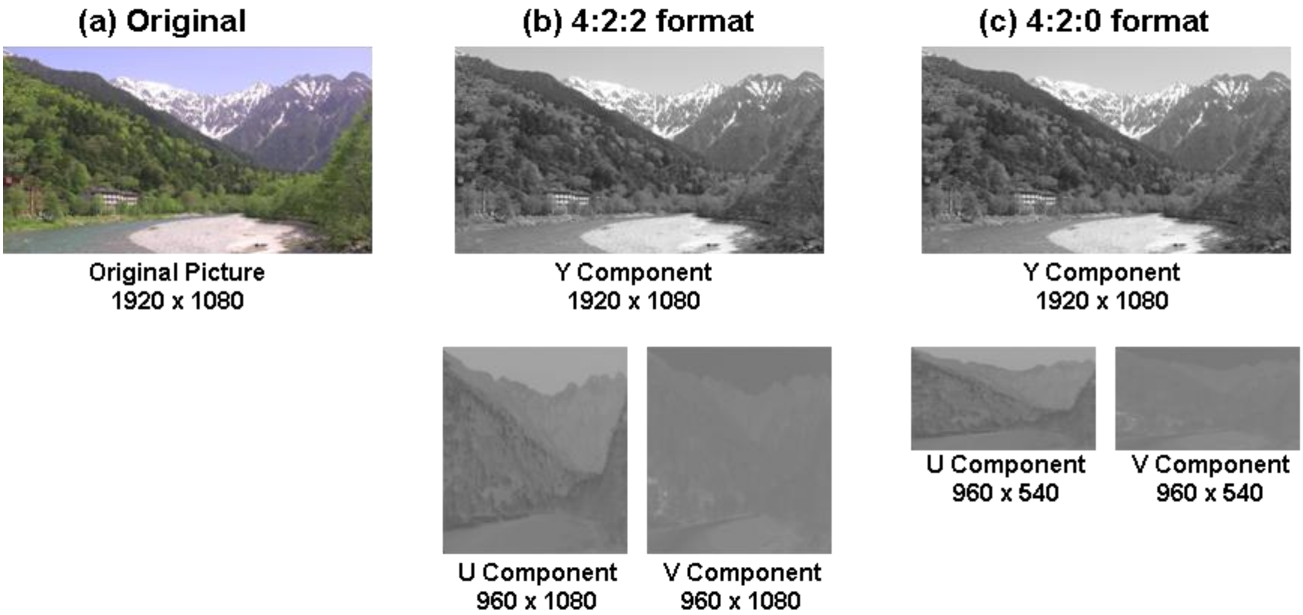



**Figure 1. 4:2:2 format and 4:2:0 format.**

### Concatenation Problem for 4:2:0 Codec

It is widely understood that for lower bit rates, 4:2:0 yields a higher picture quality, whereas for higher bit rates, 4:2:2 is better. In the professional broadcast transmission environment, 4:2:2 sampling is widely used in archiving applications such as sports contribution feeds and network distribution in which higher channel bandwidths are available. On the other hand, 4:2:0 sampling is used for SNG feeds because of limited and costly satellite bandwidth. There is, however, one major drawback for 4:2:0 chroma sampling. When two or more 4:2:0 codecs are concatenated, the chroma component gradually becomes blurred, because the codecs are usually connected via HD-SDI (a 4:2:2 interconnect), resulting in upsampling and downsampling of the chroma component.

**Requirements for Non-degraded 4:2:0**

This paragraph clarifies the requirement of the 4:2:2/4:2:0 filter set, which realizes the non-degraded 4:2:0 feature.

The first requirement is the suppression of chroma degradation after several stages of concatenation. Figure 2 shows a schematic diagram of a typical codec concatenation. In the conventional 4:2:0 codec, the chroma component gradually becomes blurred, because the codecs are usually connected via HD-SDI (a 4:2:2 interconnect), resulting in upsampling and downsampling of the chroma component. To prevent blurring, the chroma resolution of the first 4:2:0 decoder output should be preserved even after several stages of 4:2:2/4:2:0 conversion, as long as all of the codec use the same 4:2:2/4:2:0 filter set. To realize this feature, special consideration is necessary to design the 4:2:0/4:2:0 filter set.

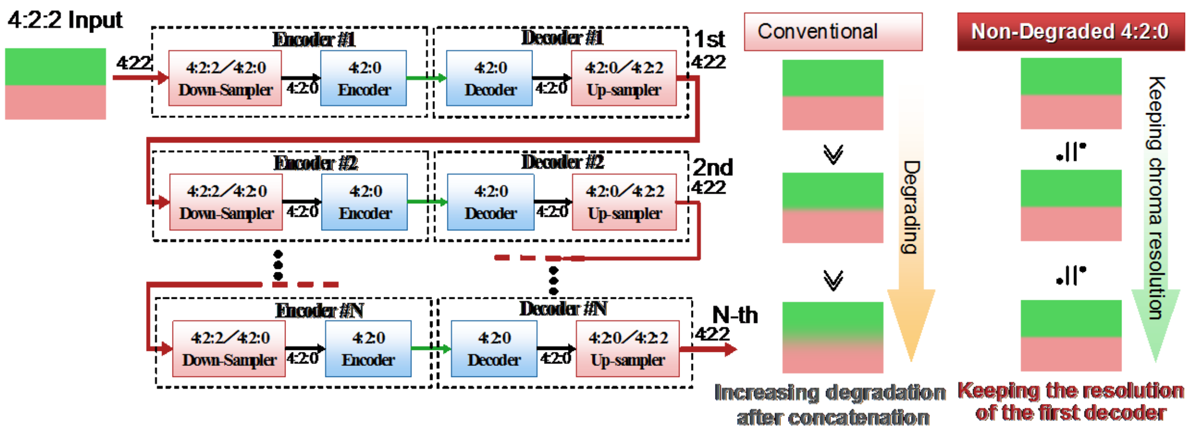



**Figure 2. Non-degraded 4:2:0 features.**

The second requirement is the compatibility with the conventional decoders. In the SNG system, there are actually several kinds of decoders. To realize the flexibility of the SNG system, the compressed stream of the non-degraded 4:2:0 encoder should be decodable at the conventional 4:2:0 decoder without any degradation or color displacement (Figure 3) This is a very important requirement.

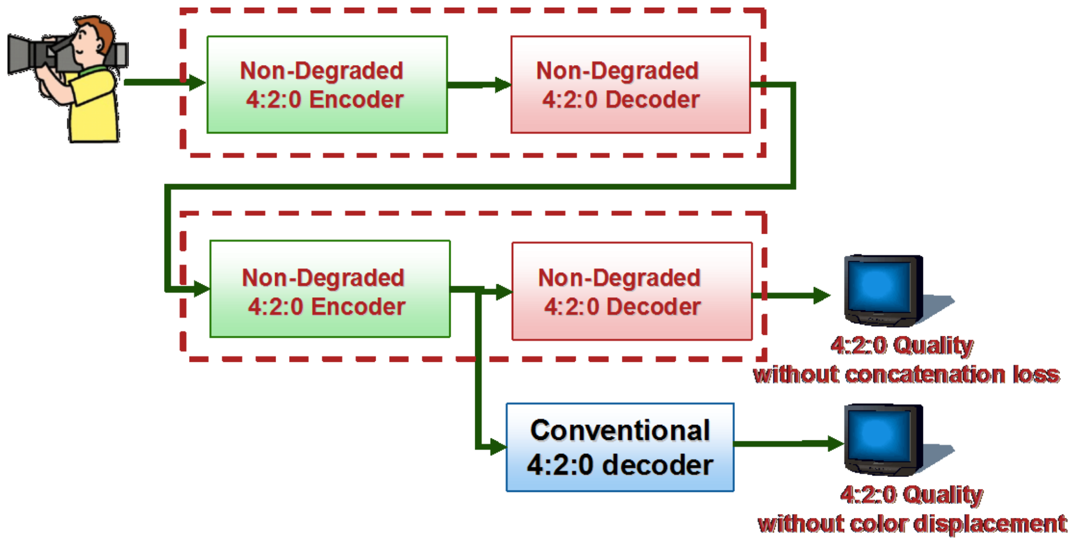



**Figure 3. Compatibility between non-degraded 4:2:0 and conventional 4:2:0 decoders.**

## Design of 4:2:2/4:2:0 Perfect Reconstruction Filter Sets for Non-degraded 4:2:0

To realize the non-degraded 4:2:0 feature, the design of 4:2:2/4:2:0 downsampling and 4:2:0/4:2:2 upsampling filter set is very important. This section explains how to design the 4:2:2/4:2:0 filter set. First, the mathematical conditions are clarified in order to realize the requirements that were set in the first section. Second, the procedure for the filter design is shown. And, finally, examples of the filter sets are shown.

### Conditions for Non-degraded 4:2:0 Feature

This section clarifies the mathematical condition to realize the non-degraded 4:2:0 feature. Figure 4 shows a schematic diagram of 4:2:2/4:2:0 and 4:2:0/4:2:2 conversions.

First, ***F0(z)*** and ***F1(z)*** are defined as z-transform of downsampling and upsampling filters, having coefficients $f0_n$ and $f1_n$ as follows, respectively.

$$F0(z) = \sum_n f0_n \bullet z^{-n}$$

$$F1(z) = \sum_n f1_n \bullet z^{-n}$$

In 4:2:2/4:2:0 conversion, the chroma signal is separated from the 4:2:2 signal, and downsampled using the lowpass filter ***F0(z)*** and the 2:1 downsampler in the vertical direction. Then, the 4:2:0 signal is formed by combining the downsampled chroma and luma.

In 4:2:0/4:2:2 conversion, the chroma signal is also separated from the 4:2:0 signal and upsampled using the 1:2 upsampler and the lowpass filter ***F1(z)*** in the vertical direction. Then, the 4:2:2 signal is formed by combining the up-sampled chroma and luma.

The following section describes the detailed condition of the 4:2:2/4:2:0 filter set to realize the non-degraded 4:2:0 feature.

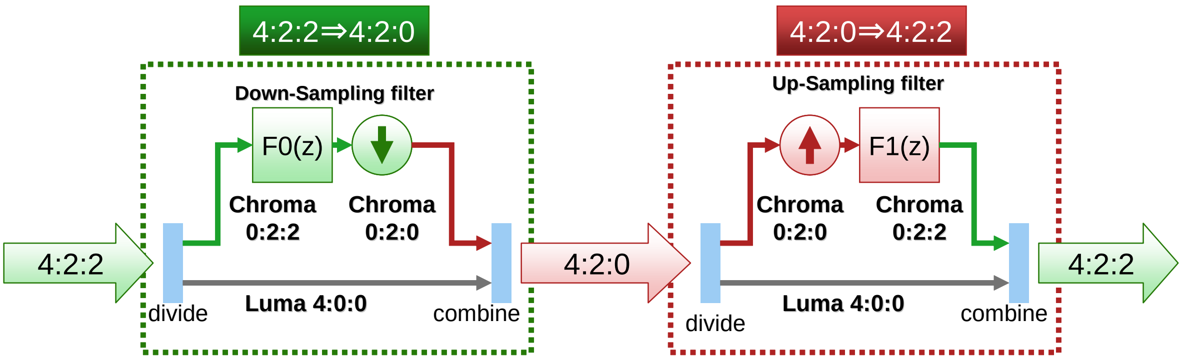



**Figure 4. Diagram of 4:2:2/4:2:0 and 4:2:0/4:2:2 conversions.**

### Perfect Reconstruction Filter Bank Condition

This paragraph clarifies the mathematical condition for the non-degraded 4:2:0 feature.

First, the 2:1 downsampling procedure is described mathematically, using Figure 5**.** The input signal is set to $a_0$, $a_1$, $a_2$, $a_3$,… and so on. Here, the 2:1 downsampler selects only the even-number signals and outputs them.

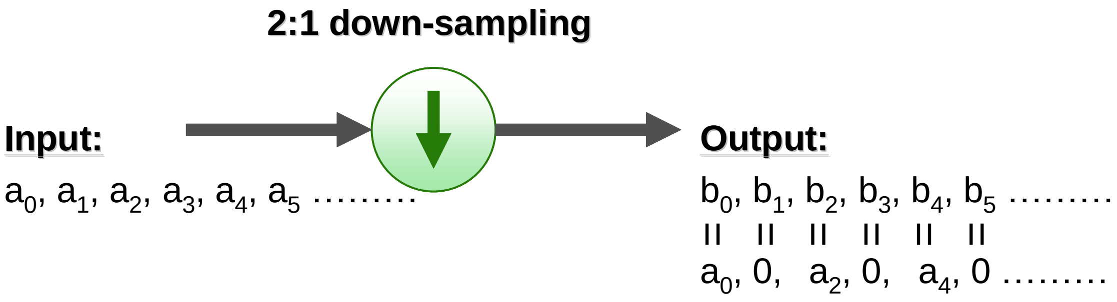



**Figure 5. 2:1 downsampling procedure.**

In this case, the output signal $\boldsymbol{b_0, b_1, b_2}$,… are described as follows:

$$b_n = \begin{cases} a_n, n = even \\ 0, n = odd \end{cases} \qquad \textbf{(Equation 1)}$$

The z-transform of input signal $\boldsymbol{\{a_n\}}$, $\boldsymbol{A(z)}$, can be described as follows:

$$A(z) = \sum_n a_n \bullet z^{-n} \quad \text{where} \quad z = e^{j\omega} \qquad \textbf{(Equation 2)}$$

By using A(z), the z-transform of output signal ***{bn}***, ***B(z)***, can be described by using ***A(z)*** as follows:

$$B(z) = \sum_n b_n \bullet z^{-n} = \sum_{n=even} a_n \bullet z^{-n} = \sum_n \left( a_n \bullet \frac{1}{2}\left(1 + (-1)^{-n}\right)\right) \bullet (z)^{-n} = \frac{1}{2}\sum_n a_n \bullet \left((z)^{-n} + (-z)^{-n}\right) =$$
$$\frac{1}{2}\left(\sum_n a_n \bullet (z)^{-n} + \sum_n a_n \bullet (-z)^{-n}\right) = \frac{1}{2} \bullet \left(A(z) + A(-z)\right)$$

**(Equation 3)**

Figure 6 shows a diagram of downsampling and upsampling filter concatenation. In Figure 6, the input signal ***X(z)*** is iteratively downsampled using filter ***F0(z)***, and upsampled using the upsampling filter ***F1(z)***. Then n-th outputs of upsampling filter, $\boldsymbol{Y_0(z), Y_1(z), \ldots Y_N(z)}$, are obtained.

By using Figure 6 and Equation 3, the condition of the non-degraded 4:2:0 feature can be derived according to the following.

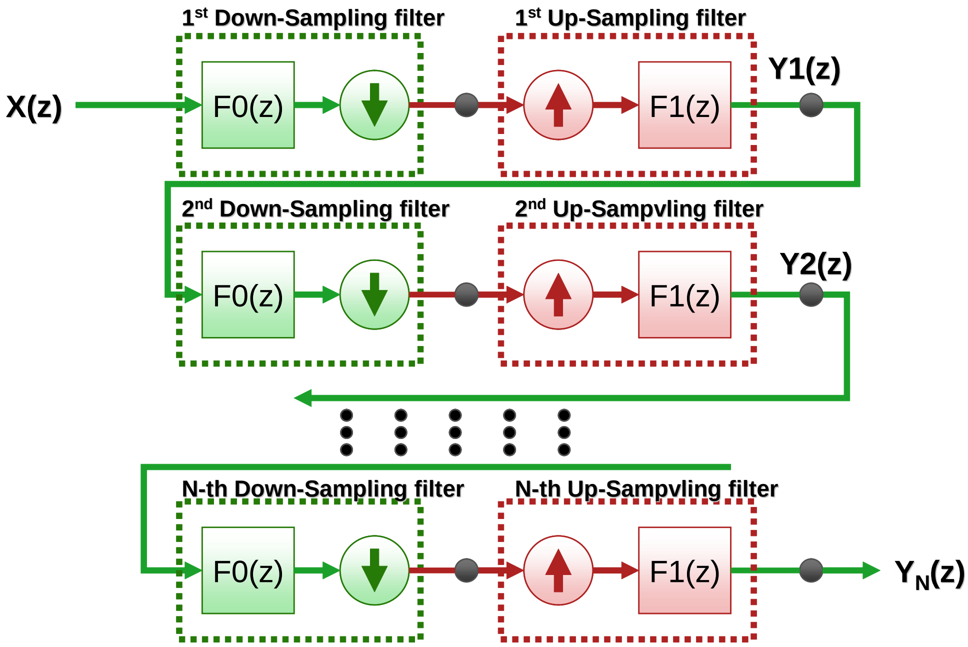



**Figure 6. Diagram of downsampling and upsampling filter concatenation.**

First, the input signal ***X(z)*** is filtered by ***F0(z)***, decimated by 2:1 downsampled, and then filtered by ***F1(z)***. By using Equation 3, $Y_1(z)$ can be expressed as Equation 4.

$$Y_1(z) = \frac{1}{2} \bullet \left( F0(z) \bullet X(z) + F0(-z) \bullet X(-z) \right) \bullet F1(z) \qquad \textbf{(Equation 4)}$$

Next, the output of the second upsampling filter, $Y_2(z)$, can be derived by applying $Y_1(z)$ to ***X(z)*** in Equation 4.

$$\begin{aligned} Y_2(z) &= \frac{1}{2} \bullet \left( F0(z) \bullet Y_1(z) + F0(-z) \bullet Y_1(-z) \right) \bullet F1(z) \\ &= \frac{1}{4} \bullet \left( F0(z) \bullet F1(z) + F0(-z) \bullet F1(-z) \right) \bullet \left( F0(z) \bullet X(z) + F0(-z) \bullet X(-z) \right) \bullet F1(z) \\ &= \frac{1}{2} \bullet \left( F0(z) \bullet F1(z) + F0(-z) \bullet F1(-z) \right) \bullet Y_1(z) \end{aligned}$$

**(Equation 5)**

The authors' goal is to keep the picture quality of the first upsampling filter output, even after the second, third, and n-th stages of downsampling/upsampling. Accordingly, the following condition must be satisfied.

$$\frac{Y_2(z)}{Y_1(z)} = \frac{1}{2} \bullet \left( F0(z) \bullet F1(z) + F0(-z) \bullet F1(-z) \right) = 1 \qquad \textbf{(Equation 6)}$$

Here, Equation 6 can be rewritten as Equation 7.

$$F0(z) \bullet F1(z) + F0(-z) \bullet F1(-z) = 2 \qquad \textbf{(Equation 7)}$$

It is noted that Equation 7 is the same condition as the "Perfect Reconstruction Filter bank", or CQF(conjugate quadrature filters) condition for the biorthogonal wavelet transform[(1)]. That is, down-sampling filter F0(z) and up-sampling filter F1(z) are equivalent to the analysis low-pass filter and synthesis low-pass filter of biorthogonal wavelet filter bank respectively.

Accordingly, $Y_3(z)=Y_2(z)$, $Y_4(z)=Y_3(z)$,..$Y_{N+1}(z)=Y_N(z)$, respectively. As a result, all of the upsampling filter outputs become mathematically the same when the Equation 7 is satisfied.

**Phase Shift Condition Conforming to 4:2:2/4:2:0 Pixel Position**

The upsampling filter and the downsampling filter should satisfy the phase shift (or group delay) conditions between 4:2:2 and 4:2:0 formats.

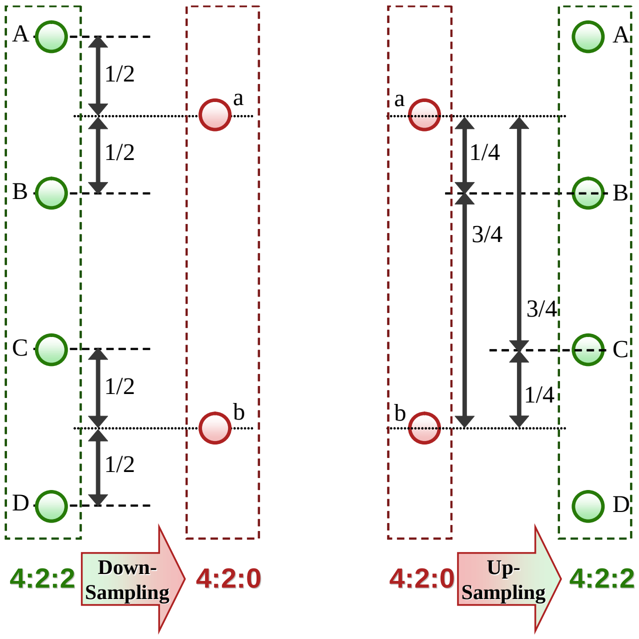



**Figure 7. Vertical shift of chroma pixels between 4:2:2 and 4:2:0 conversion for progressive.**

Figure 7 shows the vertical shift of chroma pixels between 4:2:2 and 4:2:0 for progressive. In 4:2:2/4:2:0 downsampling, the position of pixel “***a***” in 4:2:0 is 1/2 pixel below the position of pixel “***A***” in 4:2:2.

In 4:2:0/4:2:2 upsampling, two 4:2:2 pixels, “***B***” and “***C,***” are interpolated from “4:2:0” pixels “***a***” and “***b.***” Here, the position of pixel “***B***” in 4:2:2 is 1/4 pixel below the position of pixel “***a***” in 4:2:2, and the position of pixel “**C**” in 4:2:2 is 3/4 pixel below the position of pixel “***a***” in 4:2:2.

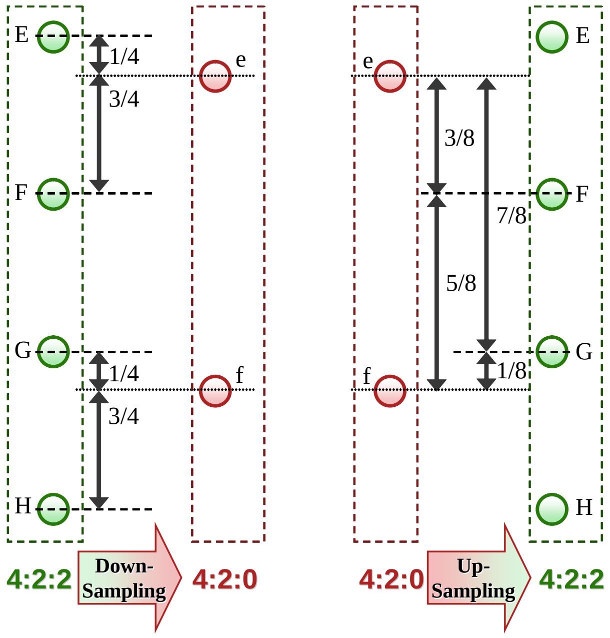



**Figure 8. Vertical shift of chroma pixels between 4:2:2 and 4:2:0 for interlace (top field).**

Figure 8 shows the vertical shift of chroma pixels between 4:2:2 and 4:2:0 for interlace (top field). In 4:2:2/4:2:0 downsampling, the position of pixel "***e***" in 4:2:0 is 1/4 pixel below the position of pixel "***E***" in 4:2:2. It is mathematically proved that only the 0 and 1/2 pixel shift can be realized using the finite length tap filter. Thus, the precise 1/4 pixel shift cannot be achieved using any kind of downsampling filter with finite length tap. Practically, the downsampling filter is designed in such a way that the pixel shift is approximately close to 1/4.

In 4:2:0/4:2:2 upsampling, two 4:2:2 pixels, "***F***" and "***G***," are interpolated from "4:2:0" pixels "***e***" and "***f***." Here, the position of pixel "***F***" in 4:2:2 is 3/8 pixel below the position of pixel "***e***" in 4:2:2, and the position of pixel "***G***" in 4:2:2 is 7/8 pixel below the position of pixel "***e***" in 4:2:2.

The phase shift characteristic of the non-degraded 4:2:0 filter should be approximately close to these values. When the compatibility is considered, the phase shift of the downsampling filter is more important.

**Example of Filter Coefficients**

In the experimental 4:2:2/4:2:0 perfect reconstruction filter set, both the downsampling filter and upsampling filter have 8-tap. This paragraph shows the designed 4:2:2/4:2:0 perfect reconstruction filter set and the design method.

**Design Method**

To design the 4:2:2/4:2:0 perfect reconstruction filter sets for non-degraded 4:2:0, it was found that both the 4:2:2/4:2:0 downsampling filter and 4:2:0/4:2:2 upsampling filter should be very carefully designed at the same time by considering the conditions described earlier.

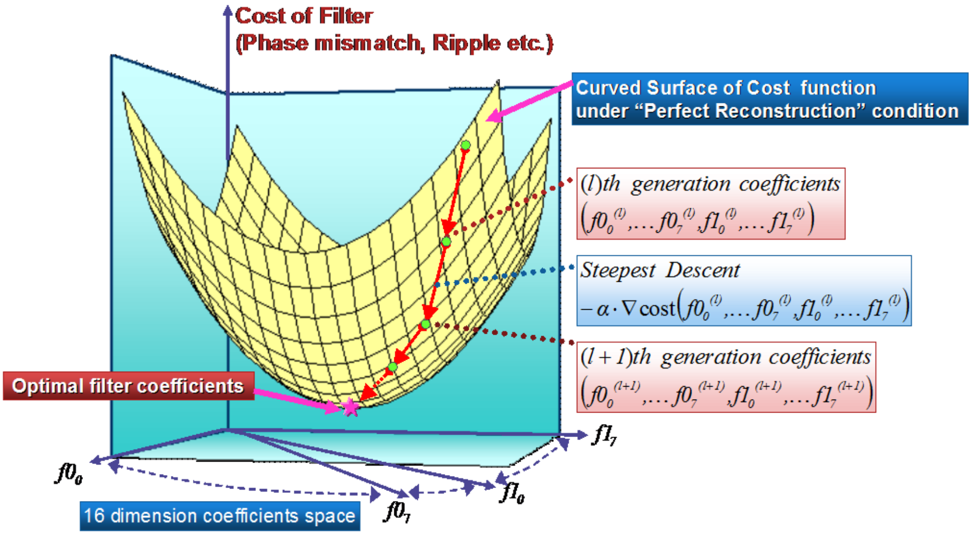



**Figure 9. Optimization of 4:2:2/4:2:0 perfect reconstruction filter sets.**

Figure 9 shows the schematic diagram of the characteristic optimization process for the design of 4:2:2/4:2:0 perfect reconstruction filter sets. Here, the coefficients of downsampling filter and upsampling filter are donated as follows:

- 4:2:2/4:2:0 downsampling filter: $f0_0$, $f0_1$, …, $f0_7$

- 4:2:0/4:2:2 upsampling filter: $f1_0$, $f1_1$, …, $f1_7$

Regarding the upsampling filter, special treatment is needed. In Figure 9, for example, coefficients $f1_0$, $f1_2$, $f1_4$, $f1_6$ are used to get the pixels such as *“**B**”* or *“**F**”* in 4:2:2, and coefficients $f1_1$, $f1_3$, $f1_5$, $f1_7$ are used to get the pixels such as *“**C**”* and *“**G**,”* respectively.

To design the optimal filter set, the following method was used.

- Definition of the Cost Function:

When we design the 4:2:2/4:2:0 perfect reconstruction filter sets, we regard the subjective quality as the most important. In order to realize high subjective quality, we use the following three evaluation criteria.

The first criterion is ringing. According to the prior studies of image coding using biorthogonal wavelets, the suppression of ringing in step response of low-pass analysis and synthesis filters is important as well as suppression of alias in frequency domain[(2)(3)(4)] because filters which considers only steep cut-off sometimes cause large ringing and degrade the subjective quality. Considering from this remark, we introduced the evaluation criterion of over-shoot in step response after down-sampling using F0(z) and up-sampling using F1(z).

The second criterion is group-delay frequency response. Especially for interlace, the group delay frequency response is more important than for progressive because both down-sampling filter F0(z) and up-sampling filter F1(0) inherently must have non-linear phase characteristics. Poor group delay characteristics of filter F0(z) and F1(z) causes jaggy especially in diagonal edges after up-sampling because the top field lines and the bottom field lines are alternatively combined in a frame. Considering from these remarks, we introduced the evaluation criteria of group-delay frequency response. In case of interlace, the following three values were used as the evaluation cost of group-delay frequency response according to the ideal group delay explained in Figure 8.

$$Cost1(\mathbf{f0}) = \int_0^{\omega_p} \left| \frac{d}{d\omega}\left( \arctan\left( \frac{\mathrm{Im}\left(\mathrm{F0}\left(e^{j\omega}\right)\right)}{\mathrm{Re}\left(\mathrm{F0}\left(e^{j\omega}\right)\right)} \right)\right) - \frac{1}{4} \right|^2 d\omega$$

$$Cost2(\mathbf{f1even}) = \int_0^{\omega_p} \left| \frac{d}{d\omega}\left( \arctan\left( \frac{\mathrm{Im}\left(\mathrm{F1}even\left(e^{j\omega}\right)\right)}{\mathrm{Re}\left(\mathrm{F1}even\left(e^{j\omega}\right)\right)} \right)\right) - \frac{7}{8} \right|^2 d\omega$$

$$Cost3(\mathbf{f1odd}) = \int_0^{\omega_p} \left| \frac{d}{d\omega}\left( \arctan\left( \frac{\mathrm{Im}\left(\mathrm{F1}odd\left(e^{j\omega}\right)\right)}{\mathrm{Re}\left(\mathrm{F1}odd\left(e^{j\omega}\right)\right)} \right)\right) - \frac{3}{8} \right|^2 d\omega$$

where,

$$F1even(z) = \sum_{n=0}^{4} f1_{2n} \bullet z^{-2n} \ , \ \ F1odd(z) = \sum_{n=0}^{4} f1_{2n+1} \bullet z^{-(2n+1)} \ ,$$

and $\omega_p$ is a pass band frequency.

The third criterion is magnitude frequency response in order to reduce alias after down-sampling and up-sampling.

At last, we defined the cost function by combining these three evaluation criteria to evaluate the performance of designed filter.

- Steepest Descent Method:

The optimal filter set is defined as a filter set that gives the minimum cost value. To get the filter set with minimum cost, the "Steepest Descent" method was used. First, the "Perfect Reconstruction Filter Bank" condition, or CQF condition for 8-tap filters can be rewritten as follows:

$$\left\{ k = -3\ldots3 \,|\, \sum_{n=0}^{7} f0_n \bullet f1_{n-2k} = 2 \bullet \delta_{0,k} \right\}$$

Assume that there are the l-th generation of filter coefficients as ($f0_0^{(l)}, f0_1^{(l)}, \ldots, f0_7^{(l)}, f1_0^{(l)}, f1_1^{(l)}, \ldots, f1_7^{(l)}$). Using these filter coefficients, the cost function can be expressed as follows:

$$\mathrm{cost}\left(f0_0^{(l)}, \ldots f0_7^{(l)}, f1_0^{(l)}, \ldots f1_7^{(l)}\right)$$

Then, the curved surface of the cost function under the condition of the "Perfect Reconstruction Filter Bank" explained earlier (Figure 9). The steepest descent direction on this curved surface can be derived as the partial differentiation of the cost function.

$$\nabla \mathrm{cost}\left(f0_0^{(l)}, \ldots f0_7^{(l)}, f1_0^{(l)}, \ldots f1_7^{(l)}\right)$$

The (l+1)-th generation's coefficients ($f0_0^{(l+1)}, f0_1^{(l+1)}, \ldots, f0_7^{(l+1)}, f1_0^{(l+1)}, f1_1^{(l)}, \ldots, f1_7^{(l)}$) are updated by updating the l-th generation's coefficients to the steepest descent direction.

$$\left(f0_7^{(l+1)}, \ldots, f0_7^{(l+1)}, f1_0^{(l+1)}, \ldots f1_7^{(l+1)}\right) =$$
$$\left(f0_7^{(l)}, \ldots, f0_7^{(l)}, f1_0^{(l)}, \ldots f1_7^{(l)}\right) - \alpha \bullet \nabla \mathrm{Cost}\left(f0_7^{(l)}, \ldots, f0_7^{(l)}, f1_0^{(l)}, \ldots f1_7^{(l)}\right)$$

where α is a step size of the iteration.

When the cost is converged to the minimum after several times of updating, the optimal filter coefficients can be regarded as optimum.

**Coefficients of "Non-degraded 4:2:0" Filter Sets for Interlace**

The design of the interlace filter set is more difficult than that of the progressive because of the nature of the nonlinear phase. Furthermore, degradation caused by inappropriate characteristics of the filter set is more sensitive in interlace format than in progressive format.

According to the procedure explained in the previous paragraph, the 4:2:2/4:2:0 perfect reconstruction filter set has been defined for the interlace format. The filter coefficients are described as follows:

**1. Downsampling filter**

The filter coefficient for interpolating 1/4 pixel position as follows:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **-0.01292** | **-0.03282** | **0.14539** | **0.57316** | **0.37762** | **-0.04072** | **-0.01008** | **0.00038** |

The pixel shift of this filter at z = 1 (frequency equal to 0) is 0.22649, i.e., 0.91/4. This is close to 1/4.

**2. Upsampling filter**

The filter coefficient for interpolating 3/8 pixel position is as follows:

| | | | |
|---|---|---|---|
| **-0.10563** | **0.82482** | **0.28434** | **-0.00352** |

The pixel shift of this filter at z = 1 is 0.38293, i.e., 3.06/8. This is almost equal to 3/8.

The filter coefficient for interpolating 7/8 pixel position is as follows:

| | | | |
|---|---|---|---|
| **0.0416** | **-0.06629** | **1.11721** | **-0.09251** |

The pixel shift of this filter at z = 1 is 0.89059, i.e., 7.12/8. This is almost equal to 7/8.

When designing the filter set, the subjective picture quality and phase and amplitude conditions were carefully considered. The proposed filter results in a good compromise between these requirements.

**Coefficients of "Non-degraded 4:2:0" Filter Sets for Progressive**

Using the same method as interlace, the 4:2:2/4:2:0 perfect reconstruction filter set for progressive format has been designed. The filter coefficients are described as follows:

**3. Downsampling filter**

The filter coefficient for interpolating 1/4 pixel position is as follows:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **-0.0025491** | **-0.01852** | **0.033479** | **0.487592** | **0.487592** | **0.033479** | **-0.01852** | **-0.00255** |

The pixel shift of this filter at z = 1 is 0.5 exactly.

**4. Upsampling filter**

The filter coefficient for interpolating 1/4 pixel position is as follows:

| | | | |
|---|---|---|---|
| **-0.13205** | **1.013644** | **0.100234** | **0.018174** |

The pixel shift of this filter at z = 1 is 0.26863393, i.e., 1.07/4. This is almost equal to 1/4. The filter coefficients for interpolation 3/4 pixel position is the upside-down version of this filter.

## Evaluation

The non-degraded 4:2:0 filter sets for interlace, which were designed to test images for performance evaluation have been applied. This section discusses the results of this performance evaluation.

### Reference Filter Set for Comparison

The non-degraded 4:2:0 filter set was compared with a conventional 4:2:2/4:2:0 filter set as a reference.

As an example of the conventional filter set for simulation, the following coefficients were used. This filter set has good cutoff frequency and group delay characteristics.

Downsampling filter:

The filter coefficient for interpolating 1/4 pixel position is as follows:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| -0.01855 | -0.0293 | 0.197266 | 0.477539 | 0.366211 | 0.048828 | -0.04004 | -0.00195 |

The pixel shift of this filter at z = 1 is 0.252916. This is almost equal to 1/4.

Upsampling filter:

The filter coefficient for interpolating 3/8 pixel position is as follows:

| | | | |
|---|---|---|---|
| -0.13281 | 0.78125 | 0.445313 | -0.09375 |

The pixel shift of this filter at z = 1 is 0.390623. This is almost equal to 3/8.

The filter coefficient for interpolating 7/8 pixel position is as follows:

| | | | |
|---|---|---|---|
| -0.02734 | 0.113281 | 0.992188 | -0.07813 |

The pixel shift of this filter at z = 1 is 0.863278. This is almost equal to 7/8.

Figure 10 shows the amplitude characteristics of the downsampling filter in each of the filter sets. Compared to the 8-tap reference filter, the cutoff of the downsampling filter (in non-degraded 4:2:0 filter set) is somewhat worse.

However, this characteristic is similar to the average of 4-tap filters. For example, the characteristic of the TM5 4:2:2/4:2:0 filter is also shown for the bottom field, that is (1, 7, 7, 1)/16. The characteristics of the average 4-tap filter and the downsampling non-degraded 4:2:0 filter are very similar. Subjective picture quality was also tested by using many different kinds of sequences, and it was found that it has not been compromised.

Considering these facts, the authors believe that the performance of the proposed filter set is sufficient to use in 4:2:2/4:2:0 conversions.

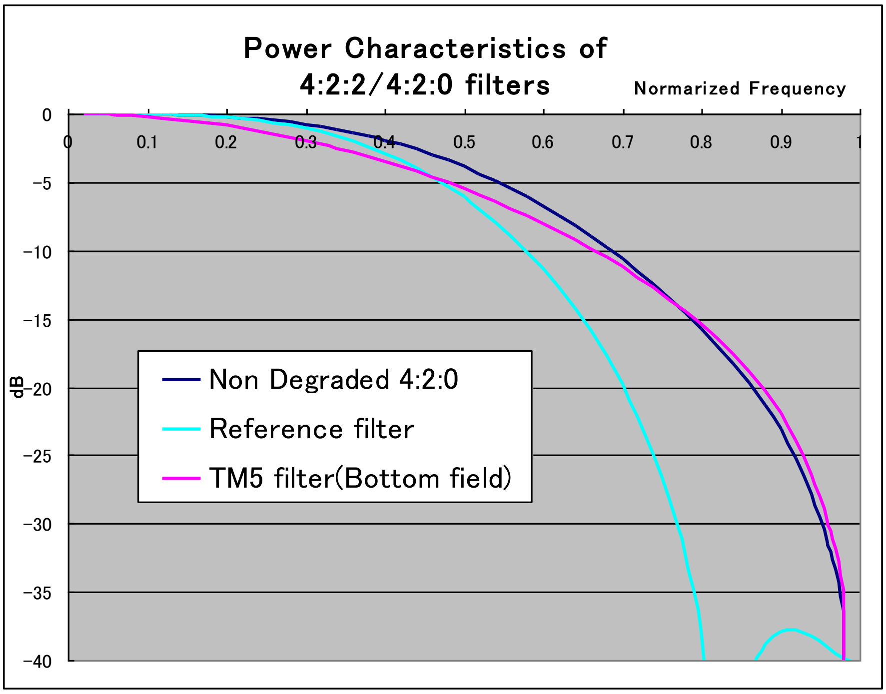



**Figure 10. Amplitude characteristics of downsampling filters.**

## Simulation Results

The PSNR and subjective picture quality were compared after several iterations of 4:2:2/4:2:0 conversion between the usual filter set and non-degraded 4:2:0 filter set.

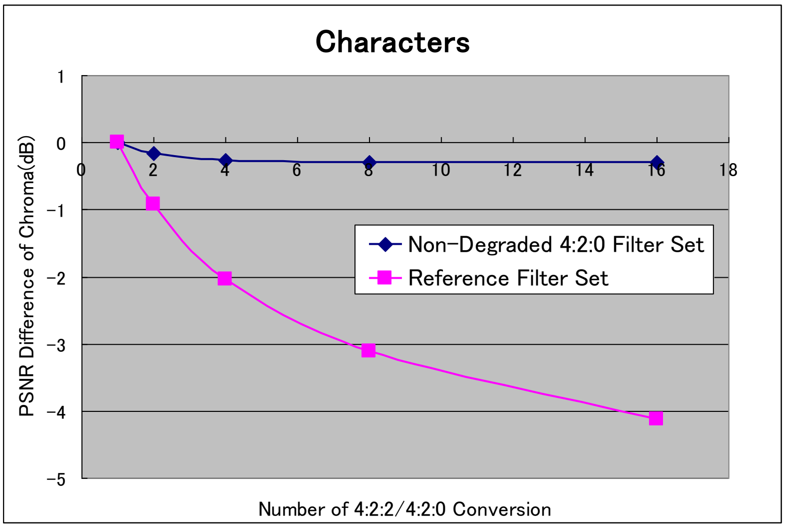
It is widely said that color degradation is more visible especially around characters or CGs that have sharper color edges after several stages concatenation of 4:2:0 codecs. Thus, a "Character" sequence (HD) made by the authors, and "F1 car" sequence (SD) made by RAI was used to show the usefulness of the non-degraded 4:2:0 filter set. The "Mobile & Calendar" sequence made by CCETT was also used as an example of a colorful sequence.

In the simulation, each of the filter coefficients is rounded to 10 bits below the decimal point because of the consideration of the implementation in the hardware.

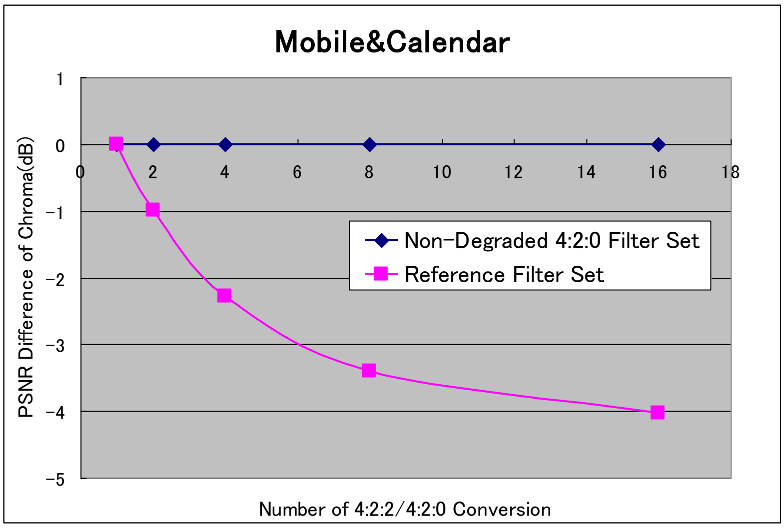

(a) "Characters" (b) "Mobile & Calendar" (c) "F1 Car"

**Figure 11. Sequences for evaluation.**

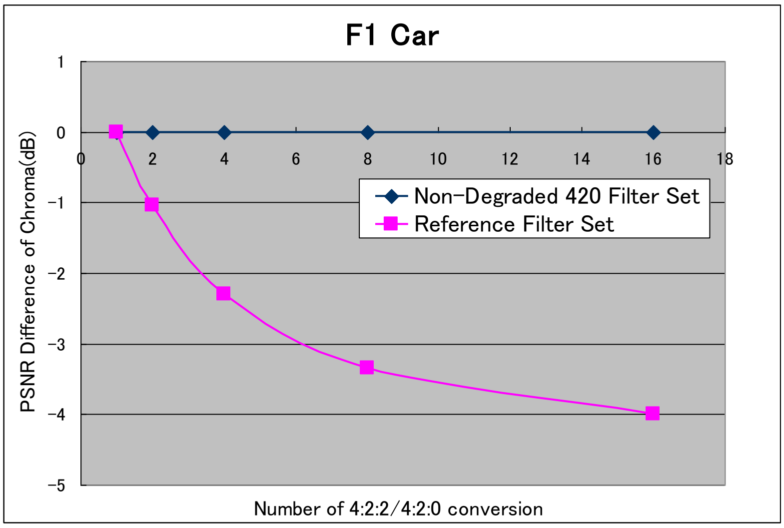
Figure 12 illustrates the PSNR difference of the chroma components (average of U and V) in the 4:2:2 format between the 1st conversion and 2, 4, 8, 16 times iterations of the 4:2:2/4:2:0 conversion. Notice that the PSNR loss of the usual filter set becomes dramatically larger. On the contrary, the PSNR of the non-degraded 4:2:0 filter, maintains almost the PSNR of the initial stages.




(a) “Characters”




(b) “Mobile & Calendar”



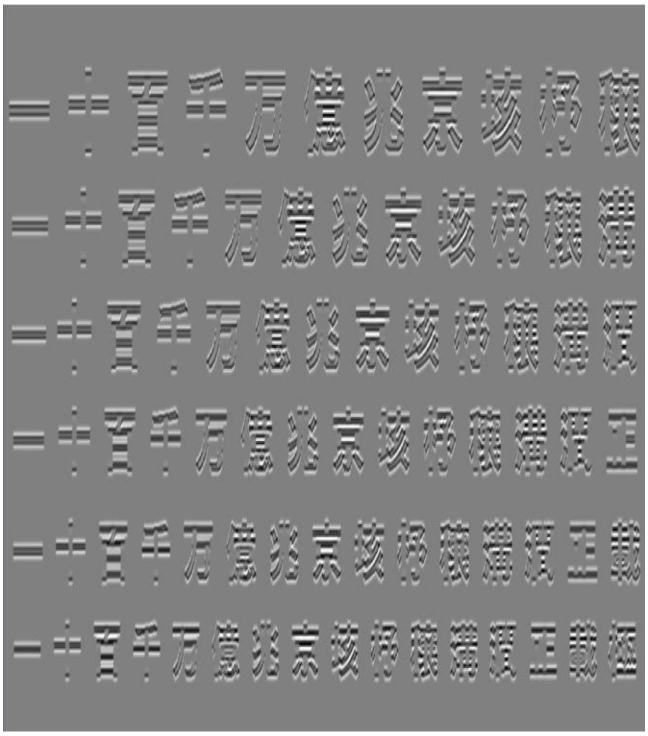

(c) “F1 Car”

**Figure 12. Chroma PSNR after 4:2:2/4:0 conversions.**

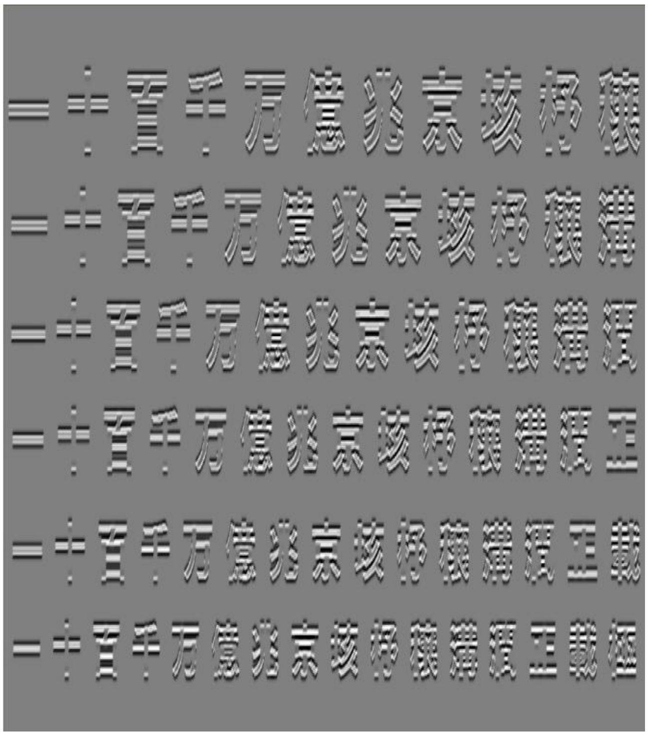
Figure 13 illustrates the difference of chroma components between the 1st and 16th 4:2:2/4:2:0 conversions. In Figure 13, the differences of U/V pixel values are multiplied by 2 in order to make the difference more visible.

When the reference filter set is used, the difference between the 1st and 16th 4:2:2/4:2:0 conversion pictures are very obvious. On the contrary, almost no difference can be seen when the non-degraded 4:2:0 filter set is used. In the “Character” sequence, a little difference can be seen in the result of the non-degraded 4:2:0 filter set, because of the rounding of the filter coefficients. However, the degree of the difference is much less than that of the reference filter sets.

Reference Filter Set

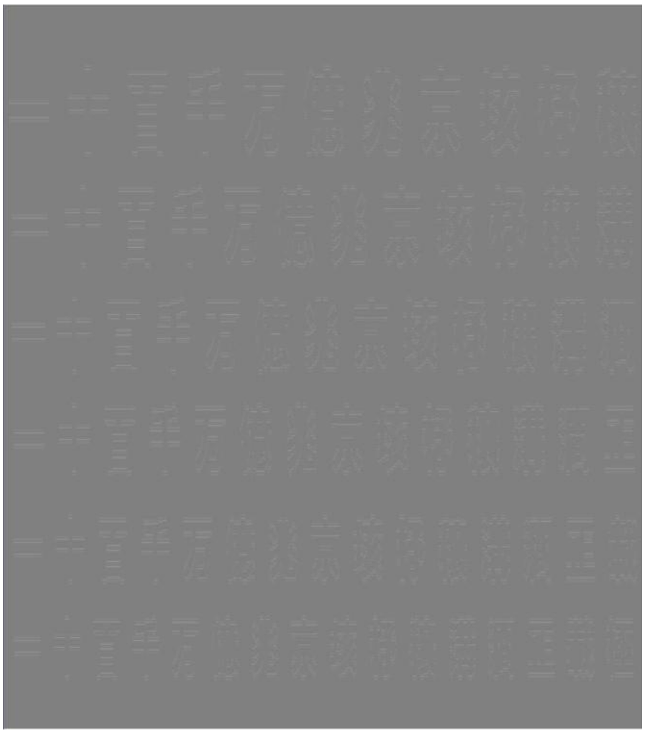


U Component



V Component

“Non-degraded 4:2:0” Filter Set



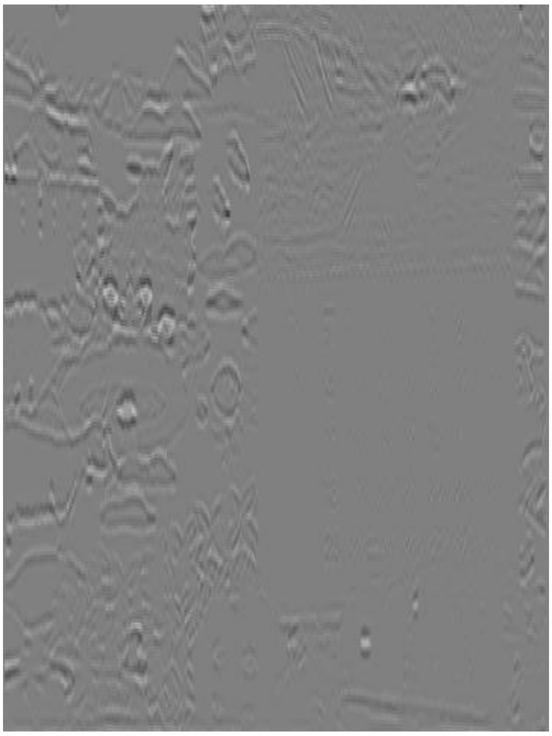
U Component

V Component

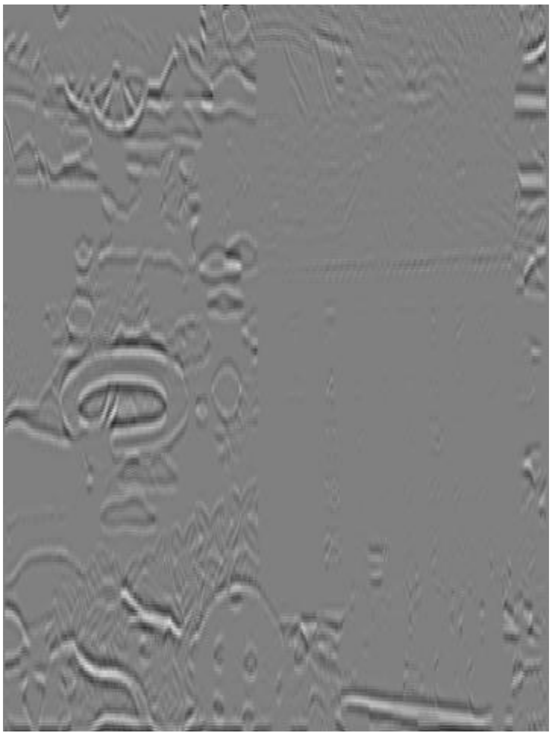
(a) “Character” Sequence

**Figure 13-1. Difference of chroma components between the 1st and 16th 4:2:2/4:2:0 conversions.**

Reference Filter Set

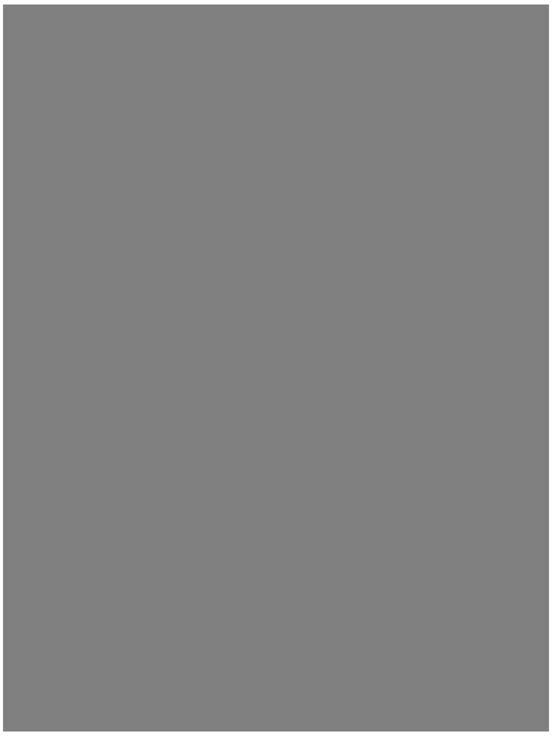


U Component

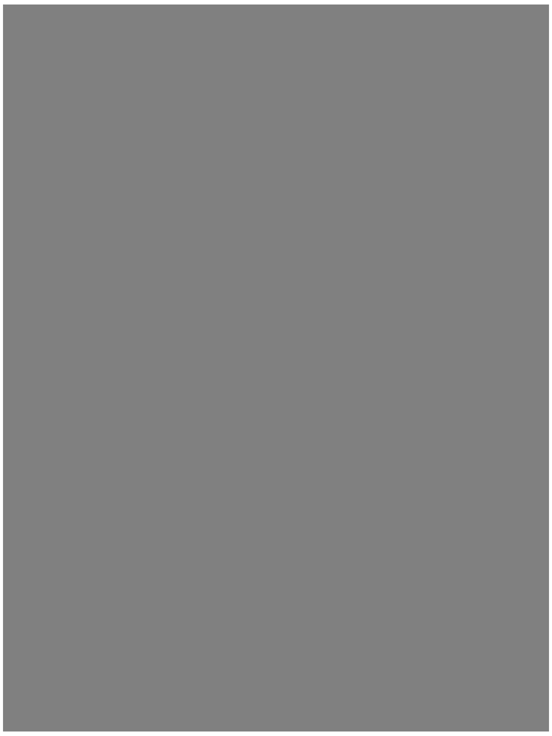


V Component

“Non-degraded 4:2:0” Filter Set



U Component



V Component

(b) “Mobile & Calendar” Sequence

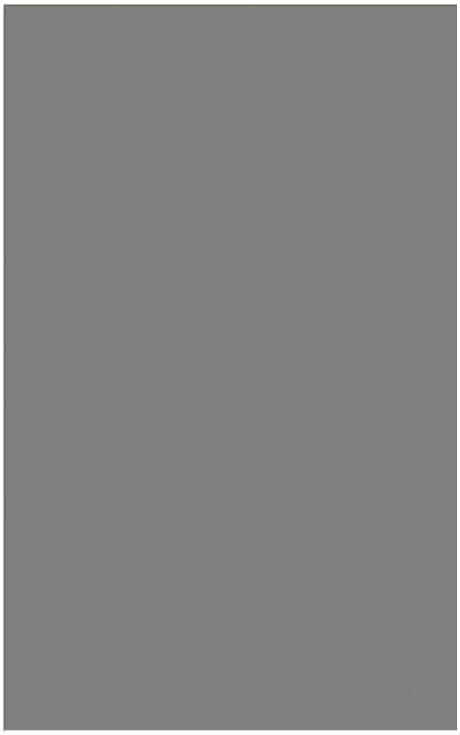
**Figure 14-2. Difference of chroma components between the 1st and 16th 4:2:2/4:2:0 conversions.**

Reference Filter Set

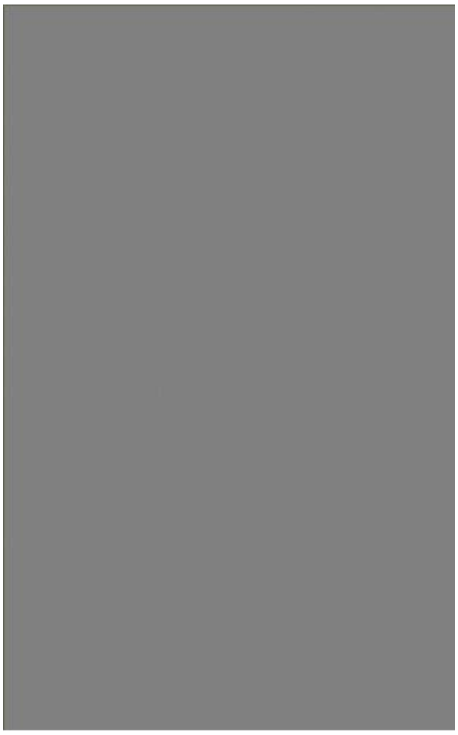
U Component

V Component

“Non-degraded 4:2:0” Filter Set



U Component



V Component

(c) “F1 Car” Sequence

**Figure 15-3. Difference of chroma components between the 1st and 16th 4:2:2/4:2:0 conversions.**

Figure 14 illustrates the area around the characters or colorful parts after several conversion iterations using the reference filter set and non-degraded 4:2:0 filter set. The numbers of 4:2:2/4:2:0 conversions are 1, 4, and 16 iterations, respectively.

When the reference filter set is used, color degradation can be seen in all of three sequences. In the “F1 Car” sequence, it is found that the area around the numbers becomes increasingly yellow with each conversion. Similarly, in the “Character” sequence, the area inside the characters becomes increasingly blurry. In the “Mobile & Calendar” sequence, the red and green colors expand to the white background.

On the contrary, the converted picture using the non-degraded 4:2:0 filter set maintains the initial stage’s quality even after 16 conversions in these three sequences.

Of course, the degree of blurring depends on the filter set. However, as long as the filter set does not satisfy the non-degraded 4:2:0 conditions, the converted picture becomes more-or-less blurred.

Reference Filter Set

1 times 4 times 16 times

“Non-Degraded 4:2:0” filter Set

1 times 4 times 16 times

(a) “Character” Sequence

**Figure 16-1. Comparison of converted pictures after several stages.**

Reference Filter Set

1 times 4 times 16 times

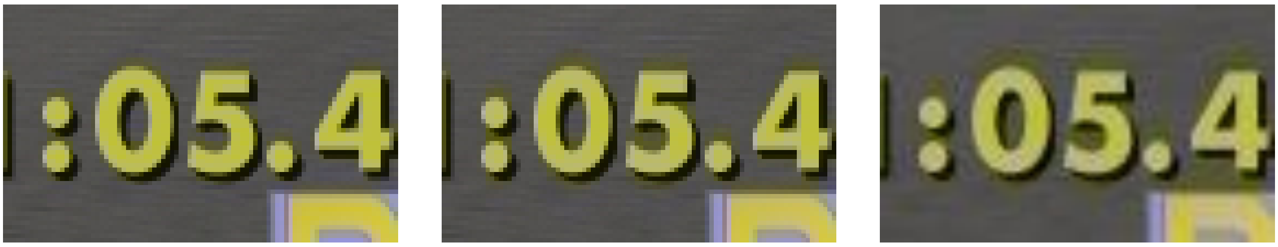
“Non-Degraded 4:2:0” filter Set

1 times 4 times 16 times

(b) “Mobile & Calendar” Sequence

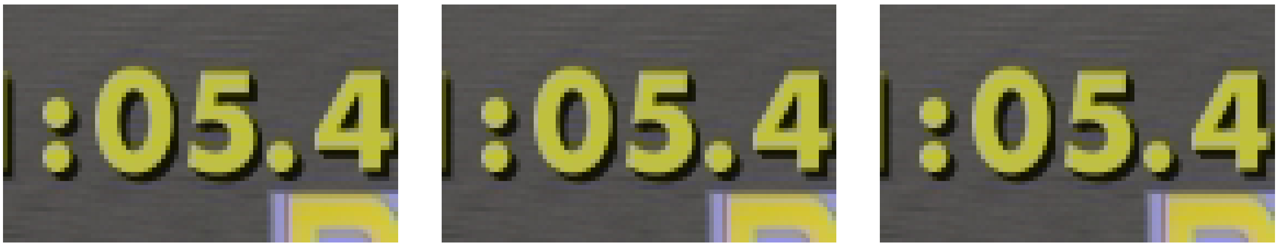
**Figure 17-2. Comparison of converted pictures after several stages.**

Reference Filter Set



1 times | 4 times | 16 times

"Non-Degraded 4:2:0" filter Set



1 times | 4 times | 16 times

(c) "F1 Car" Sequence

**Figure 18-3. Comparison of converted pictures after several stages.**

Figure 15 shows the part of pictures after 4:2:2/4:2:0 downsampling and 4:2:0/4:2:2 upsampling, using three kinds of filter combinations. The conditions are as follows:

(1) Both the downsampling filter and upsampling filter are those of the reference filter set.

(2) Both the downsampling filter and upsampling filter are those of the non-degraded 4:2:0 filter set.

(3) The downsampling filter is that of the non-degraded 4:2:0 filter, and the upsampling filter is that of the reference filter. This is simply the case in Figure 3 in which a bit stream produced by an encoder with the non-degraded 4:2:0 downsampling filter is decoded and displayed by the decoder with the reference upsampling filter.

These results illustrate that these three pictures are almost identical. In case (2), ringing and jaggy around edge area are preferably suppressed. That is, our non-degraded 4:2:0 filter can realize the almost equivalent subjective quality as the reference filter set. Further more, there is no pixel shift or degradation even in case (3). We have extensively evaluated the converted pictures of case (1), case (2) and (3) using various kinds of sequences in hardware, and we concluded that this filter set has good characteristic enough to apply to professional use.

As a result, a very flexible codec arrangement can be realized in SNG/microwave link application.


(1) Reference filter set

(2) Non-degraded 4:2:0 filter set

(3) Non-degraded 4:2:0 downsample filter + Reference upsampling Filter

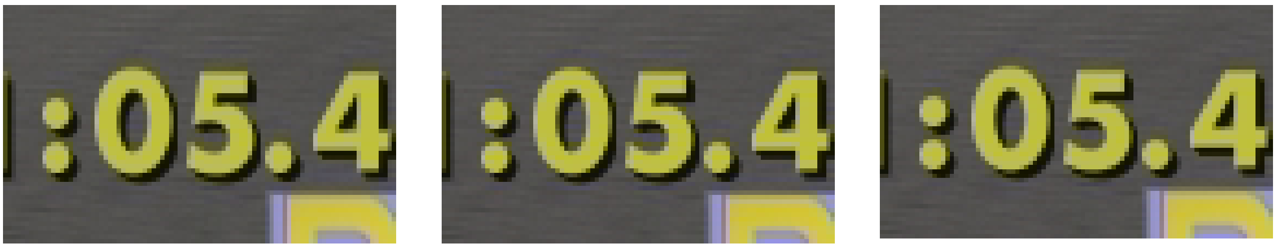
(a) "Character" Sequence

(1) Reference filter set

(2) Non-degraded 4:2:0 filter set

(3) Non-degraded 4:2:0 downsample filter + Reference upsampling Filter

(b) "Mobile & Calendar" Sequence




(1) Reference filter set

(2) Non-degraded 4:2:0 filter set

(3) Non-degraded 4:2:0 downsample filter + Reference upsampling Filter

(c) "F1 Car" Sequence

**Figure 15. Comparison of converted pictures after several stages.**

### Implementation in the Hardware

The designed filter coefficients in the 4:2:0 MPEG-4 AVC codec have also been implemented. It is found that the color degradation is also imperceptible, except for the coding error caused by MPEG-4 AVC.

In order to communicate the use of this filter set, VANC information can be used. Regarding the detail design, further study is needed.

## Conclusion

When two or more 4:2:0 codecs are concatenated, the chroma components gradually become blurred because the codecs are usually connected via HD-SDI (a 4:2:2 interconnect), resulting in upsampling and downsampling of the chroma component.

We have explained an example of a new chroma downsampling and upsampling filter set for use with 4:2:0 codecs whereby chroma degradation is completely suppressed after several stages of concatenation. It is shown that almost no degradation can be seen up to 16 times stages of concatenation. Furthermore, this filter set also has a compatibility with the conventional decoders. That is, the compressed stream of non-degraded 4:2:0 encoder can be decodable at the conventional 4:2:0 decoder without any degradation or color displacement.

In conclusion, the 4:2:2/4:2:0 perfect reconstruction filter set is very useful to the SNG/ENG application.

## Reference

(1) Ingrid Daubechies, “Ten Lectures on Wavelets”, 1992, Society for Industrial & Applied

(2) W. Aasno, T. Takeuchi, M. Ikehara, “Filterbanks for Less Ringing Artifacts”, IEICE Transactions (Japanese Edition), VOL. J84-A, No.5, pp.576-585, 2001

(3) M.A.M. Rodrigues, Eduardo A. B. da Silva, Paulo S. R. Diniz, “MINIMIZING RINGING EFFECT ON IMAGES CODED AT LOW BIT RATES WITH WAVELETS”, ICIP98, vol.3, pp.668-672, 1998.

(4) John D. Villasenor, Benjamin Belzer, Judy Liao , “Wavelet Filter Evaluation for Image Compression”, IEEE TRANS. on Image Processing, Vol. 4, No.8, pp.1053-1060, August 1995