SYSTEM-ON-SILICON SOLUTION FOR HIGH QUALITY CONSUMER VIDEO PROCESSING-
THE NEXT GENERATION

Markus Schu, Dirk Wendel, Christian Tuschen, Marko Hahn, Ulrich Langenkamp
Micronas Munich GmbH, P.O.Box 800949, D-81609 Munich, Germany
e-mail: <first>.<last>@micronas.com

Abstract — This paper presents the next generation of a
highly integrated low cost single-chip video processing
system-on-chip with outstanding feature content and
performance. The IC\(^1\) unifies the complete processing
chain between tuner output (CVBS) and RGB
processing. It comprises an improved motion adaptive
algorithm for high quality de-interlacing and up-
conversion including a special cinematic source
processing. Implemented are picture-in-picture, split
screen and manifold picture manipulation and
improvement capabilities as well as color decoding and
all necessary A/D and D/A conversions. All field
memories are realized on-chip and store a picture in
4:2:2 format up to SVGA resolution. The IC meets
economical demands with increased functionality.

1 INTRODUCTION

Image sequences occur at various picture rates. Source
material normally exists in 24p, 25p, 30p, 50i or 60i Hz
format (i-interlaced, p-progressive), whereas video is
usually shot and broadcast at 50i and 60i Hz. For flicker-
free display of 50i (60i) Hz video, up-conversion to 100i
(60p) Hz is desirable. Using sophisticated motion
compensation techniques \([1,4,5]\), the sample rate
conversion leads to a good motion portrayal in the
displayed material. However, this technique is up to now
still expensive for consumer application and often reveals
shortcomings in line flicker reduction performance and
resolution, especially in still areas of images. The
 economical implementation of a cost effective scan rate
conversion including a very good line flicker reduction \([7]\)
leads to a sophisticated combination of non motion
compensated up-conversion methods with automatic
treatment of video and film sources.

To meet the market’s demand for flexible multi-channel
displays, the chip includes 2 independent color decoders,
several additional inputs, and two parallel processing
chains, which allow the manipulation of two video signals.
Each signal passes a noise reduction, can be scaled to any
format (by linear or panoramic expansion or compression
in horizontal and vertical direction), is up-converted and
improved by several algorithms. Both channels are
processed and improved independently.

The result is another simplification of the standard TV set
configuration, like illustrated in a functional block diagram
in figure 1.

\[\text{Figure 1: Typical configuration for a standard TV set}\]

2 SYSTEM OVERVIEW

The architecture of the IC consists of three main blocks:
front end processing, dual channel processing including
embedded memory and display processing (figure 2).

The following features are combined in this architecture:
• Analog CVBS, Y/C, RGB and YCrCb inputs
• Support of one digital ITU656 input
• Input sources: 24p, 25p, 30p, 50p, 60p, 50i, 60i
• 2 multi-standard color decoders incl. 4H comb filter
• Flexible motion adaptive de-interlacing/up-conversion
• High performance 3D motion detection
• Horizontal and vertical compression and expansion
  including 5 segment panorama
• 8 Mbit embedded memory for field stores
• Application modes for picture-in-picture, split screen
  and multi-picture displays, PC based displays
• 3D spatial-temporal picture enhancement technique
• Adaptive luminance peaking, LTI and CTI

\(^1\) The IC is commercially available as VSP9407.
To utilize the dual channel processing for flexible display arrangements, different application modes can be selected. Available are motion adaptive full-screen-mode (FSM), split-screen-mode (SSC), multi-picture-mode (MUP), snapshot-mode (SPS) and PC-mode (PCF). Each mode optimizes the internal configuration of the IC to enable the desired display features.

Known from previous ICs [1] are FSM-mode (including a PiP insertion), SSC-mode (double-window) and MUP-mode (tuner-scan, two live pictures). When selecting the SPS mode (Figure 4), the current picture is stored in the memory, whereas the running TV program is still shown. It is possible to show either the stored picture (e.g. including information like phone numbers) or the running TV program or both together. For example this can be a PiP insertion as reminder for the stored picture.

The PiP insertion does not necessarily need to be decimated in size to show the complete picture content. It can also display only a non-scaled part of the content. This new interpretation of PiP is usable to show a stock-ticker from one channel inside the main channel (Figure 5).

PC-modes can be used to show a VGA or SVGA signals on the screen. Either full-screen PC displays, PC split-screen applications (e.g. PC next to TV) or PiP insertions (TV on PC, PC on TV) are possible. The horizontal and vertical frequency of PC signals as well as the signal polarities are detected automatically.
3 MOTION ADAPTIVE UP-CONVERSION AND DE-INTERLACING

The complete double scan conversion algorithm is specified and well suited for creating a de-interlaced progressive scan output (double horizontal and single vertical scan frequency, 50p/60p Hz) as well as for up-converted double horizontal and double vertical frequency output (100i/120i Hz). Below the fundamental principle is explained for the up-conversion to 100i/120i Hz.

An economical way to increase the picture rate from 50i/60i Hz to 100i/120i Hz is to repeat fields as long as their contents are available. This algorithm – called field repetition – only needs one field memory and is therefore very cost attractive. Looking to the input field, the line scan pattern (α/β) and the displayed motion phase (A/B/C/D/S) are linked:

| motion: | A^α | B^β | C^α | D^β |
| still: | S^α | S^β | S^α | S^β |

Using the field repetition this connection remains unchanged. So repeating one field forces also a repetition of the line scan pattern. The resulting sequence is

| motion: | A^α | A^α | B^β | B^β | C^α | C^α | D^β | D^β |
| still: | S^α | S^α | S^β | S^β | S^α | S^α | S^β | S^β |

Doubling the field rate to 100i/120i Hz removes the large area flicker completely. Nevertheless the line scan pattern and motion phase change frequency remain at 50i/60i Hz, so line flicker is still visible in still areas. Although not reduced it is not visible for moving objects because flicker artifacts are covered by the moving objects.

In figure 6 (top) a one-dimensional display is illustrated. An object (represented by a square box) is displayed with and without motion. For field repetition (top) with motion it is clearly visible that the object is displaced from the true motion line (dotted) in every second output field. A viewer tracks the moving object, realizes the actual object position at the displaced position but expects it on the true motion position. Now the viewer perceives two objects, at the real displaced and the expected true position. This artifact is known as “double contour” and can only be avoided by using motion compensation techniques [1]. Nevertheless for certain kinds of source material (camera pan, material containing moving objects) this simple scan rate conversion method leads to good results.

Adding a second field memory enables a frame repetition. This algorithm also doubles the change frequency of the line scan pattern and motion phase. Thus the line flicker is removed, too. The output sequence looks like

| motion: | A^α | B^β | A^α | B^β | C^α | D^β | C^α | D^β |
| still: | S^α | S^β | S^α | S^β | S^α | S^β | S^α | S^β |

Figure 6: Scan rate conversion methods
Annoying is the step back on motioned objects (figure 6, mid) which leads in extreme motion judder. This artifact disqualifies the algorithm for displaying motion. Otherwise no judder can occur on completely still pictures. Here the best possible display quality can be reached without any flicker.

For the major part of source material both situations occur. Areas of the sequence are in motion, others are still. The field repetition would cause line flicker artifacts in still areas whereas the frame repetition results in motion judder in areas of motion. To combine the advantages of both algorithms – field and frame repetition – a third algorithm is applied. Because the line scan pattern can only switch between fields but not between areas of the image, the frame repetition method is used. Only a new motion display method must be designed. This looks like:

\[
\text{mot. area: } A^\alpha A^\beta B^\alpha B^\beta C^\alpha C^\beta D^\alpha D^\beta \ldots \\
\text{still area: } S^\alpha S^\beta S^\alpha S^\beta S^\alpha S^\beta S^\alpha S^\beta \ldots 
\]

The motion phases A, B, C, D are modified to A', B', C', D' [7] to fit the new line scan pattern. Now a pixelwise blend between the still area and motion area algorithm (soft-mix) can be realized (figure 6, bottom).

To control this switch the IC contains a high performance 3D motion detection circuit, which analyzes the picture concerning moving objects. As a result, an information is assigned to each pixel, indicating if the pixel is still or in motion. Figure 7a shows typical source material, where line flicker reduction (frame repetition) should be applied for the still parts and the modified field repetition method should be used for the moving parts. In Figure 7b the result of the motion detection circuit is visualized. The marked areas are in motion.

![Figure 7: a) source material for soft blend mode](image1.png)

Adding detectors which indicate high amount of motion (global motion detector) or no motion (global still detector) allow to switch between the three different methods. Like illustrated in Figure 8 the global motion detector controls the switch-over between the soft blend method and the field repetition. The global still detector controls the transition between soft blend and frame repetition.

The global still detection is robust in case of noisy pictures.

![Figure 8: Flexible up-conversion concept](image2.png)

If the input is a film source (PAL 2-2 pull down, NTSC 3-2 pull down) special output motion sequences can be generated to increase the output picture quality. For example, a field jam algorithm for NTSC film material is implemented. On field jam always two fields with different line scan pattern but the same motion phase are combined to an output frame to preserve the frame resolution. To handle these input sources a 3\(\text{rd}\) dimension was added to the algorithm (Figure 8) which allows separate parameter sets for film sources. The switch is controlled by a film mode detector which provides information about the type of the input source (Video, PAL film or NTSC film) and the film phase (2 for PAL film, 5 for NTSC film).

To meet the customer’s need the implemented double scan conversion can be programmed very flexibly. Motion sequences as well as line scan pattern sequences, interpolation types and reactions to detectors can be set in a wide range. In addition all detection units can be adjusted separately.

### 4 A CHOICE OF NEW ALGORITHMS

The new chip combines known picture processing functionality with new developed algorithms.

#### 4.1 Picture improvements

Separating noise from the images simultaneously with sharpening these images is one key idea. Thereby, the different frequency channels of the video signal are processed in different ways in order to achieve the most attractive effect on the human visual system. New algorithms were developed including luminance transition improvement (LTI) and contrast adaptive peaking for luminance component, as well as chrominance transition improvement (CTI) for the color components. The LTI preprocesses the luminance signal to obtain a better result for the following contrast adaptive peaking. The adaptive peaking as well as the chrominance and luminance transition improvement (CTI, LTI) are described in detail in [6].
4.2 Noise measurement for picture content

To improve the quality of noisy pictures it is advisable to apply some sort of noise reduction. As these algorithms usually decrease the quality of pictures with little noise, it is highly desirable to apply a noise adaptive mechanism, which makes strong corrections in pictures with poor quality, and small corrections in pictures with good quality. To control this mechanism, it is necessary to measure the extent of noise in the picture. The standard for this measurement is the Peak Signal to Noise Ratio (PSNR). However, as it needs a reference-picture it is not possible to use it for TV-applications, where such reference is not available.

While in former algorithms the noise was measured in the blanking interval [1], by filling this part with additional information, now it is necessary to do the noise measurement within the active image.

![Diagram of noise measurement algorithm](image)

Figure 9: New noise measurement algorithm

Furthermore transmission of the picture content is going digital so the noise is only in the picture content. This results in the problem of separating the noise from the content of the picture, and to develop a figure which is sensitive to noise, but does not vary too much with different structures in the wide range of possible images. White noise makes a contribution to all frequencies in the picture. Therefore, it should be possible to detect it in the high frequencies of the image.

![Figure 10: a) source picture, b) detected homogenous areas](image)

Figure 10: a) source picture, b) detected homogenous areas

It is necessary to detect areas that are homogenous and quantify the high frequencies only in these areas. This works because it is unlikely to only have fine structure in the picture.

4.3 Scaling units and panorama mode

To adapt the input picture to the used display, the integrated up-sampling unit allows linear and even non-linear picture expansion, commonly known as ‘panorama mode’.

The picture width and height are user-adjustable in steps of two output pixels and one line. For panorama mode the picture is divided into 5 segments – horizontally and vertically. For every segment the characteristics of the expansion can be defined separately.

![Figure 11: a) source picture, b) example of horizontal panoramic distortion, c) example of vertical panoramic distortion, d) example of horizontal and vertical panoramic distortion](image)

4.4 Special Application Modes

The IC provides complex feature content. It allows attractive arrangements for the display of different input sources and the transition between one display mode to another. In the already described MUP mode it is possible to move the picture to be updated over the background to its desired position.
It is possible to blend dynamically from one input source to another (e.g. programme change) or use the blend function statically in the case of a PiP insertion.

For so called ‘violence protection’ or applications with conditional access, a mosaic mode is applicable to scramble the picture.

5 ARCHITECTURE

The economical implementation of the algorithms described above on a single IC including the necessary field and line memories requires an embedded DRAM technology. To preserve pin, software and feature compatibility with current and future TV concepts and to reduce the time-to-market, the development of standardized video processing cores was essential [2, 3].

The key characteristics of the IC are listed in Table 1. The field and line memories are realized on-chip. In addition ADC and DAC cores are included. Table 2 lists several on-chip cores. Figure 15 presents the layout of the described chip.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Transistor count</td>
<td>16 Millions</td>
</tr>
<tr>
<td>Data clocks [MHz]</td>
<td>20.25, 40.5, 27, 36, 72</td>
</tr>
<tr>
<td>Supply voltages</td>
<td>1.8V, 3.3V</td>
</tr>
<tr>
<td>Dissipation</td>
<td>1200 mW</td>
</tr>
<tr>
<td>Package</td>
<td>P-MQFP 80</td>
</tr>
</tbody>
</table>

Table 1: Key characteristics of the new IC

<table>
<thead>
<tr>
<th>Core functions</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field memories</td>
<td>8 x 1 Mbit</td>
</tr>
<tr>
<td>4xADC (RGB, fast-blank)</td>
<td>8 bit @ 40.5 MHz</td>
</tr>
<tr>
<td>2xADC (CVBS, Y/C)</td>
<td>9 bit @ 20.25 MHz</td>
</tr>
<tr>
<td>3xDAC (YUV, RGB)</td>
<td>9 bit @ 81 MHz</td>
</tr>
<tr>
<td>C800 microcontroller</td>
<td>8051 compatible</td>
</tr>
</tbody>
</table>

Table 2: On-chip cores

Figure 12: MUP with moving PiP

Figure 13: a) b) input, c) d) result of blend function

Figure 14: a) source picture, b) example for mosaic distortion

Figure 15: Layout of the new IC
Various versions of the described chip-family are available [2] or are under development which cover the whole range from 50/60 Hz interlace (single-scan) up to high-end with motion compensation. They are hardware compatible to allow the use of one chassis-concept for different types of TV-sets.

![Figure 16: Versions](image)

### 6 CONCLUSION

Due to increased need for cost effective functionality and performance in the consumer TV market this new IC has been designed. The IC is able to process two channels in parallel, enabling manifold picture arrangements and animations. The flexible upconversion concept gives line-flicker reduction without annoying motion judder and treats film-sources accordingly. This new IC has low system cost. It is another member of the hardware compatible family for flicker-free TV.

### 7 ACKNOWLEDGEMENTS

The team of authors wish to thank all people involved in this project. We are especially grateful for the contribution of Peter Rieder and Günter Scheffler to this article.

### 8 REFERENCES


MARKUS SCHU was born in St. Wendel, Germany, on May 14, 1967. He received the M.Sc. degree from University of Saarland in 1992. In June 1992 he joined the Panasonic European Research and Development Center. In July 1995 he joined the Siemens AG, Semiconductor Group, in Munich, now Infineon Technologies. Since October 2000 the consumer division has been a part of Micronas semiconductors. There he has been working in the field of scan rate conversion, noise reduction, image enhancement, motion detection, motion estimation, letter box detection and sample rate conversion. He was the responsible concept manager for the described IC. He received the first prize in the 2000 ICCE outstanding paper award for the best poster presentation [1].

MARKO HAHN was born in Rüdersdorf, Germany in September 1970. He received his Diploma of Electrical Engineering from the University of Ilmenau and joined the Siemens AG, Semiconductor Group in 1998 that became Infineon Technologies in 1999. He has worked as concept engineer for Infineon’s consumer division and he is with Micronas, now. He has been working in the field of sample and scan rate conversion, noise measurement, picture improvement and on chip bus interface concepts.

DIRK WENDEL was born in Dortmund, Germany in July 1970. He received his Diploma of Electrical Engineering in 1996 from the University of Dortmund and joined Infineon Technologies (formerly Siemens) in 1996. He worked as concept engineer for picture-in-picture (PiP) ICs. Now he is with Micronas working on high-integrated TV and multimedia products. Beside his work on the overall concept, he has been specifically responsible for the frontend-processing of the described IC.

CHRISTIAN TUSCHEN was born in Dortmund, Germany in 1969. He received his M.Sc. degree in Electrical Engineering from the University of Dortmund in 1996. In 1996 he joined the Siemens AG, Semiconductor Group, now Infineon Technologies. In 2000 the consumer division became part of Micronas Semiconductors. As system concept engineer he worked on several video IC projects, focusing on the development of video enhancement algorithms. Currently he is responsible for scan rate conversion and de-interlacing, motion estimation and motion compensation. He was the first place winner in the 2000 ICCE outstanding paper award for the best poster presentation [1].

CHRISTIAN TUSCHEN was born in Dortmund, Germany in 1969. He received his M.Sc. degree in Electrical Engineering from the University of Dortmund in 1996. In 1996 he joined the Siemens AG, Semiconductor Group, now Infineon Technologies. In 2000 the consumer division became part of Micronas Semiconductors. As system concept engineer he worked on several video IC projects, focusing on the development of video enhancement algorithms. Currently he is responsible for scan rate conversion and de-interlacing, motion estimation and motion compensation. He was the first place winner in the 2000 ICCE outstanding paper award for the best poster presentation [1].

ULRICH LANGENKAMP was born in 1952 in Wuppertal, Germany. He received his Diploma of Electrical Engineering in 1976 from the University of Bochum. As design engineer and later on as program manager he was involved in the development of the first digital TV chipset at ITT and many subsequent products for TV applications at ITT, Siemens and Infineon. Since 1989 his focus has been on embedded DRAM products.