Overview

In video compression algorithms, the quantization parameter (QP) is usually selected based on the relative complexity of the region in the picture as well as the overall bit usage. However, complexity-based rate-control algorithms do not take into account the fact that more complex objects, such as human faces, are more sensitive to degradation during perceptual video compression. To improve the overall perceived quality of the image, it is important to classify human faces as regions of interest (ROI) and preserve as much detail in those regions as possible. The challenge is developing a reliable algorithm that will operate in real time. This white paper details a low-complexity solution that is able to run on a single-core digital signal processor (DSP) as part of an encoder implementation.

Skin-tone macroblock detection

The proposed solution is a low-complexity, color-based skin-tone detection which classifies skin-tone macroblocks (MBs) as ROI MBs and non-skin-tone macroblocks as non-ROI MBs. An MB can be defined as a 16×16 block of pixels. The classification of ROI MBs and non-ROI MBs is based on empirical thresholds applied to the mean of the color components. These threshold values were defined after extensive research using material that covers various races. According to this classification and a modified rate control (RC) that smoothly assigns different levels of quality, we can increase visual quality (VQ) in human faces. The new RC assigns a lower QP to ROI areas compared to non-ROI areas while maintaining the overall bits-per-frame budget.

Erosion and dilation

Erosion and dilation algorithms are used to refine detection — reduce false positives and missed MBs. These morphology algorithms use classified neighbors’ information to fill holes (missed MBs) and locate isolated blocks (false positives). False positive ROI MBs lead to flawed allocation of important bits, while missed ROI MBs create a rough region perception.

Erosion helps to find false positives and mark them as non-ROI. Dilation, on the other hand, finds holes in skin regions (like eyes or mouth in faces regions) and marks them as ROI. In Figure 1, areas in pink show detected MBs as ROI. The face on the left shows detected skin

Figure 1. MBs detected as ROI in pink. Left image: ROI without applying morphology algorithms. Right image: ROI detection using morphology algorithms.
regions without applying morphology algorithms, and as a result, the eyes are not part of the ROI. Alternatively, the figure on the right shows the complete face is marked as ROI, as morphology algorithms have been applied.

Erosion and dilation algorithms can be implemented in pre-processing or merged inside the encoder. When used in pre-processing, all the MBs of a frame are classified as an ROI or non-ROI before being encoded. All neighboring MBs’ skin information is used to decide if a MB skin classification is a false positive, a hole, or if it is correct, and accordingly make the correct ROI or non-ROI classification.

When the erosion and dilatation algorithms are merged inside the encoder, only top, left, top-left and top-right skin MB neighbors’ information are available for making refinement decisions, but this version is suitable for low-latency applications. Figure 2 shows in pink detected MBs as ROI. The figure on the left shows results for a ROI detection merged inside the encoder. The figure on the right uses all MBs neighbors’ information in order to return a final ROI classification.

Activity gradient threshold

In addition to the erosion and dilation algorithms, an MB activity gradient threshold is implemented to reduce the amount of false positives, especially when videos have many small faces, such as faces in a crowd. Background faces in a crowd are not given ROI treatment.

An implementation based on 8×8 pixel blocks detection for Luma and 4×4 pixel blocks for Chroma components (in case of 4:2:0 video format) improves the algorithm precision compared to 16×16 pixel block detection. If two or more of the four blocks from a MB are classified as skin blocks, then the complete MB is marked as skin MB.

Pre-processing decision

It’s also important to implement another pre-processing step inside the ROI algorithm in order to eliminate frames that have too many MBs marked as ROI, in which case it is pointless to do frame bit redistribution. If more than 30 percent of a frame is detected as skin, all the MBs are remarked as non-ROI.
Decimation process

Finally, a decimation process will reduce processing cycles and increase channel density per core. This process skips some pixel values in order to get the mean of the color components blocks. For Luma components, rather than getting the mean value of 64 8-bit pixels, we decimate in steps of four and get the mean of only four 8-bit pixels per 8×8 blocks. For Chroma 4×4 components blocks, the decimation step used is two. Decimation decreases ROI classification precision. However, it results in a fair tradeoff when preprocessing multiple HD channels in a single core.

Figure 3 and Figure 4 show a VQ comparison of using or not using ROI detection as part of H.264 encoder. Figure 3 was encoded at a low bit rate to stress the VQ difference. This sequence has many small water droplets in movement. Small moving objects drain rate control’s bit budget, resulting in a visual degradation of the face when ROI RC modification is not applied.

Figure 3. Chromakey sequence, H.264 encoded. Left image: No ROI applied. Right image: ROI applied.

Figure 4 shows content similar to that found in video conferencing, where the background is usually static and faces are the critical information to transmit.

Figure 4. Video conference content. Left image: No ROI RC is applied. Right image: ROI RC is applied.
Video market trends

The current video market demands a low-complexity implementation of skin-tone detection algorithms with a highly accurate classification. ROI detection can be implemented in a video frame pre-processing stage, or with less accuracy, it can be merged inside a standard video codec. A low-complexity implementation gives the advantage of fast decision-making (fewer cycles) when determining if an MB is part of an ROI area. Fast decisions with a low-complexity classification algorithm allow a real-time ROI detection implementation on low-power processors for high-channel-density scenarios. Additionally, a low-complexity ROI implementation allow encoders improve overall video quality in applications including video broadcast, video conference, video security and smart cameras.

TI technology for skin-tone macroblock detection

The ROI detection algorithm was implemented and tested on Texas Instruments’ (TI’s) 1-GHz TMS320C674x floating-point DSP on TI’s DaVinci™ TMS320DM816x video processor. The solution is XDAIS compliant. Initially, no optimized implementation permits ROI pre-processing of three channels 1080p60. After using DSP-specialized MAC and SIMD instructions, performance was boosted to six channels of 1080p60. In video processors with three video accelerators (IVAHD), like TI’s DM816x video processor, it is possible to encode three HD channels using ROI information and still have room on the DSP for more pre-processing, such as audio or text-detection algorithms.

Figure 5 shows data and control flow implemented on TI’s DM816x video processor. From a data point of view, skin-tone detection gets YUV input buffers and generates a MB map, which is appended to the YUV data and used as metadata information for the encoder. From a control point of view, the integrated ARM® Cortex™-A8 runs the main application, which invokes process calls on the C674x DSP with raw input data.
available in DDR. The C674x DSP generates ROI information in DDR and informs the ARM® Cortex™-A8 when the frame is done; then the ARM Cortex-A8 invokes a process call on the ARM Cortex-M3, and ROI information is padded as metadata to the Cortex-M3. Once frame encoding is done, the ARM Cortex-M3 informs the ARM Cortex-A8, and the described process starts again.

As part of the video transcoder (VTC) demo on TI’s DaVinci™ DM816x video processor, developers have the option to mark detected ROI MBs with white dots for a real-time verification of the ROI detection algorithm. An example of this capability can be seen in Figure 6. The VTC demo also includes the option of encoder clips with ROI detected regions using TI’s IVAHD H.264 encoder. Using ROI detection MB map, TI’s IVAHD H.264 encoder Rate Control (RC) smoothly redistributes frame bits between ROI and non-ROI regions for a better video quality perception.

To improve the overall perceived quality of images that focus on human faces, skin-tone macroblock detection offers a low-complexity solution. This paper has shown how different techniques can be implemented in order to increase accuracy and maintain HD channel density for real-time pre-processing of the ROI detection algorithm on TI’s DSPs, including the C674x DSP on TI’s DaVinci DM816x video processor. TI’s VTC demo offers the option of ROI H.264 encoding and a real-time visualization of ROI detection.

For more information about TI’s VTC demo, please visit: [www.ti.com/truviewvtcdemo](http://www.ti.com/truviewvtcdemo)

For more information about TI’s DaVinci DM816x video processor, please visit: [www.ti.com/dm8168](http://www.ti.com/dm8168)

![Figure 6. TI’s VTC demo with ROI detection visualization.](image-url)

**Conclusion**

To improve the overall perceived quality of images that focus on human faces, skin-tone macroblock detection offers a low-complexity solution. This paper has shown how different techniques can be implemented in order to increase accuracy and maintain HD channel density for real-time pre-processing of the ROI detection algorithm on TI’s DSPs, including the C674x DSP on TI’s DaVinci DM816x video processor. TI’s VTC demo offers the option of ROI H.264 encoding and a real-time visualization of ROI detection.

For more information about TI’s VTC demo, please visit: [www.ti.com/truviewvtcdemo](http://www.ti.com/truviewvtcdemo)

For more information about TI’s DaVinci DM816x video processor, please visit: [www.ti.com/dm8168](http://www.ti.com/dm8168)
Texas Instruments Incorporated and its subsidiaries (TI) reserve the right to make corrections, modifications, enhancements, improvements, and other changes to its products and services at any time and to discontinue any product or service without notice. Customers should obtain the latest relevant information before placing orders and should verify that such information is current and complete. All products are sold subject to TI's terms and conditions of sale supplied at the time of order acknowledgment.

TI warrants performance of its hardware products to the specifications applicable at the time of sale in accordance with TI's standard warranty. Testing and other quality control techniques are used to the extent TI deems necessary to support this warranty. Except where mandated by government requirements, testing of all parameters of each product is not necessarily performed.

TI assumes no liability for applications assistance or customer product design. Customers are responsible for their products and applications using TI components. To minimize the risks associated with customer products and applications, customers should provide adequate design and operating safeguards.

TI does not warrant or represent that any license, either express or implied, is granted under any TI patent right, copyright, mask work right, or other TI intellectual property right relating to any combination, machine, or process in which TI products or services are used. Information published by TI regarding third-party products or services does not constitute a license from TI to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property of the third party, or a license from TI under the patents or other intellectual property of TI.

Reproduction of TI information in TI data books or data sheets is permissible only if reproduction is without alteration and is accompanied by all associated warranties, conditions, limitations, and notices. Reproduction of this information with alteration is an unfair and deceptive business practice. TI is not responsible or liable for such altered documentation. Information of third parties may be subject to additional restrictions.

Resale of TI products or services with statements different from or beyond the parameters stated by TI for that product or service voids all express and any implied warranties of the associated TI product or service and is an unfair and deceptive business practice. TI is not responsible or liable for any such statements.

TI products are neither designed nor intended for use in military/aerospace applications or environments unless the TI products are specifically designated by TI as military-grade or "enhanced plastic." Only products designated by TI as military-grade meet military specifications. Buyers represent that they have all necessary expertise in the safety and regulatory ramifications of their applications, and acknowledge and agree that they are solely responsible for all legal, regulatory and safety-related requirements concerning their products and any use of TI products in such safety-critical applications, notwithstanding any applications-related information or support that may be provided by TI. Further, Buyers must fully indemnify TI and its representatives against any damages arising out of the use of TI products in such safety-critical applications.

TI products are not authorized for use in safety-critical applications (such as life support) where a failure of the TI product would reasonably be expected to cause severe personal injury or death, unless officers of the parties have executed an agreement specifically governing such use. Buyers represent that they have all necessary expertise in the safety and regulatory ramifications of their applications, and acknowledge and agree that they are solely responsible for all legal, regulatory and safety-related requirements concerning their products and any use of TI products in such safety-critical applications. Buyers must contact TI for available human machine interface (HMI) safety studies, if applicable.

TI products are neither designed nor intended for use in automotive applications or environments unless the specific TI products are designated by TI as compliant with ISO/TS 16949 requirements. Buyers acknowledge and agree that, if they use any non-designated products in automotive applications, TI will not be responsible for any failure to meet such requirements.

Following are URLs where you can obtain information on other Texas Instruments products and application solutions:

<table>
<thead>
<tr>
<th>Products</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio</td>
<td>Automotive and Transportation</td>
</tr>
<tr>
<td>Amplifiers</td>
<td>Communications and Telecom</td>
</tr>
<tr>
<td>Data Converters</td>
<td>Computers and Peripherals</td>
</tr>
<tr>
<td>DLP® Products</td>
<td>Consumer Electronics</td>
</tr>
<tr>
<td>DSP</td>
<td>Energy and Lighting</td>
</tr>
<tr>
<td>Clocks and Timers</td>
<td>Industrial</td>
</tr>
<tr>
<td>Interface</td>
<td>Medical</td>
</tr>
<tr>
<td>Logic</td>
<td>Security</td>
</tr>
<tr>
<td>Power Mgmt</td>
<td>Space, Avionics and Defense</td>
</tr>
<tr>
<td>Microcontrollers</td>
<td>Video and Imaging</td>
</tr>
<tr>
<td>RFID</td>
<td></td>
</tr>
<tr>
<td>OMAP Mobile Processors</td>
<td></td>
</tr>
<tr>
<td>Wireless Connectivity</td>
<td></td>
</tr>
</tbody>
</table>

TI E2E Community Home Page: e2e.ti.com

Mailing Address: Texas Instruments, Post Office Box 655303, Dallas, Texas 75265
Copyright © 2012, Texas Instruments Incorporated