57.1: Invited Paper: Key Requirements for High Quality Picture-Rate Conversion

Claus Nico Cordes
NXP Semiconductors, Research, Eindhoven, The Netherlands

Gerard de Haan
Philips Research Laboratories, Eindhoven, The Netherlands

Abstract
Past LCD-TV generations suffered from a poor motion portrayal, causing the blurring of moving objects. Hence, various techniques have been implemented to improve their motion portrayal, of which the widespread introduction of motion compensated picture-rate conversion in TV systems is an essential part. However, a careful design of such algorithms is critical, as otherwise very annoying artifacts may decrease their value. In this paper, we will give an overview of the key requirements for high quality motion-compensated picture-rate conversion, as implemented in state-of-the-art system-on-chips, and visually illustrate the impact of individual measures on picture quality.

Keywords: motion blur, motion judder, motion compensation, picture-rate conversion, frame-rate conversion

1. Introduction
The perceived picture quality of the early LCD-TV generations was far from perfect. There are two characteristics of the Liquid Crystal (LC) display that can hold responsible for the observation that particularly the motion portrayal was far inferior to that of the earlier CRT-TV [1]:

- The slow response time of the LC material
- The picture sampling & hold of the LC display

Over time, the slow response has been decreased considerably by panel makers through material innovations, while the effect has been further reduced by display processing (overdrive). The negative consequences of the picture sampling & hold have been largely eliminated by increasing the display picture-rate from 50/60Hz to 100/120Hz, and more recently even to 200/240Hz.

After tackling the major causes of poor motion portrayal of the LCD, the flaws of source and processing emerge as the next big challenge. Since video content is mostly captured at relatively low picture rates, like 24Hz for film- or 50/60Hz for video-cameras, high quality Picture-Rate Conversion (PRC) is essential to profit at all from the fast-responding 240Hz LCD-panels.

It is broadly recognized, see e.g. [2], that the eye-tracking of the human viewer forces PRC algorithm designers to apply advanced Motion Estimation (ME) and Compensation (MC) techniques. Simpler PRC methods using repetition or linear interpolation of images cause motion blur and/or judder which renders the advances in display technology useless.

Advanced MC-PRC algorithms, which in theory enable perfect motion portrayal on modern LCDs, have already been introduced in CRT-TVs in 1995 [3]. In practice, however, MC-PRC is an art more than a science, by which we imply that any known algorithm will occasionally fail. Clearly, the main challenge in designing a MC-PRC algorithm is to limit the frequency of the occurrence of such “failures”, and particularly to avoid perceptually annoying artifacts.

In this paper, we shall discuss the key requirements and lessons-learned in the design of state-of-the-art MC-PRC algorithms, as implemented in e.g. the TV550 [4] from NXP Semiconductors.

In Section 2, the various ingredients will be illustrated by showing their visual impact from an end-user perspective, while we draw our conclusions in Section 3.

2. Key requirements of MC-PRC
Since the introduction of MC-PRC in TV systems, display technology has improved considerably. Modern displays exhibit high spatial resolutions, high contrast ratios and large screen sizes. As a consequence, the quality requirements for MC-PRC have increased over time, as the visibility of even small processing artifacts on these improved displays was guaranteed.

This has pushed the algorithm designer further on the learning curve. Looking back, we recognize the following aspects of high quality MC-PRC as most important:

- True motion estimation
- Robust interpolation
- Occlusion handling
- Global fallback
- Large velocity range
- Robust film mode detection

All above ingredients are necessary to yield a sufficiently high picture quality for mainstream and high-end TV segments. In the following sections, each of these aspects is elaborated and illustrated.

2.1 True motion estimation
Many motion estimation algorithms have been proposed in literature, yet only the category of so-called true-motion estimation algorithms can be used for MC-PRC. The specific aim of these algorithms is to determine the ’true’ motion of the scene, i.e. the 2D motion generated by the projection of a 3D ’real world’ movement onto a series of 2D images.

As a consequence, algorithms based only on residue-optimization are generally unsuitable, as they generate vector fields that are optimal in terms of match error, yet contain many errors in areas with repetitive detail. Figure 1 illustrates the output of a basic algorithm within this category, i.e. Full Search (FS), next to the output of an algorithm that is optimized for true-motion estimation, i.e. 3DRS [5].

It is apparent that FS generates a significantly noisier vector field than 3DRS, which results in clearly visible errors in the interpolation, as illustrated in Figure 2 (left).
To avoid such errors, algorithms aimed at true motion estimation typically incorporate a smoothness term in the minimization process, taking into account the fact that the vector with the lowest residue does not necessarily correspond with the true/correct motion vector. The example of 3DRS shows that this results in clearly improved interpolation quality.

Next to smoothness, the accuracy of motion vectors is also important. This accuracy is typically expressed by the quantization level of motion vectors, e.g. 1 (integer) or ½ pixel accuracy. Our experience has shown that sub-pixel accuracy has clear advantages over 1-pixel integer accuracy for two reasons. Most obviously, limited vector accuracy can result in errors of half the quantization level (e.g. ½ pixel errors in the case of ¼ pixel accuracy). This can result in a loss of spatial resolution and detail flicker. More troublesome, however, is that entirely wrong vectors may be found in repetitive structures. This effect, which is shown in Figure 3, has some resemblance to artifacts in FS vector fields. In practice, an accuracy of ¼ pixel (achieved by e.g. bilinear interpolation) largely avoids such artifacts.

Finally, the spatial resolution of the vector field is important. Ideally, we’d like to estimate the correct motion vector for every pixel of the image. However, pixel-wise motion estimation is not only computationally expensive; it also increases sensitivity to errors in repetitive structures, as smaller image parts are less uniquely defined within the image.

By jointly estimating the motion of multiple pixels (typically in blocks), this problem is greatly reduced. However, this generally introduces the problem of a lower spatial resolution, as illustrated in Figure 4 (left) with the example of blocks of 32x32 pixels. Clearly, larger blocks prevent the vector field from closely following object boundaries, often resulting in artifacts.

To achieve a trade-off between both effects, either a compromise is made within the block size, e.g. using 8x8 blocks for SD/HD resolutions, or the resolution of the block matching differs from the block grid; here, larger blocks are used to calculate the match error (e.g. 16x16), yet vectors are assigned to a smaller (e.g. 8x8) central part of the matched block.

2.2 Robust interpolation

Given true motion vectors, the subsequent interpolation may seem to be simply the case of averaging motion compensated pixels. However, motion estimation algorithms may fail for various reasons, such as:

- Multiple objects fall into a single block
- The motion is complex, e.g. rotation or zoom
- The motion is ill-defined, e.g. transparency or occlusion

Hence, we cannot assume that motion vectors are always correct, as otherwise highly visible artifacts may occur. This means that additional robustness measures are required in the interpolation part. For this, many methods exist, that typically have in common that they are based on the mismatch between both motion compensated pixels, and the mismatch between both no-motion compensated pixels.

The underlying assumption is that if the mismatch of the motion compensated pixels is high, either in absolute terms or relative to the non-motion compensated pixels, the motion vector may have been unreliable. In such cases, a local fallback is used, which is often the non-motion compensated average.

Figure 1 Motion vector overlay (in pseudo colors) of (left) full search, and (right) true motion estimation using 3DRS.

Figure 2 Interpolation results using (left) full search block matching, and (right) true motion estimation using 3DRS.

Figure 3 Motion vector overlay of 3DRS using (left) 1-pixel, and (right) ¼ pixel accuracy. The object is a rotating wheel.

Figure 4 Motion vector overlay of 3DRS using blocks of (left) 32x32 pixels, and (right) 8x8 pixels on a 720x576 input image.

Figure 5 Interpolation results using (top) motion compensated average, and (bottom) cascaded median.
There are multiple reasons for this choice, one of them being that the non-motion compensated average results in a temporal averaging of the images, which is a perceptually unobtrusive artifact. Also, interpolation errors are most visible at the borders of non-moving areas (such as subtitles or logos), and the above choice results in a natural bias towards the correct interpolation of such areas. Figures 5 and 6 illustrate this improvement with the example of the so-called cascaded median [6].

Figure 6 Interpolation results using (left) motion compensated average, and (right) cascaded median.

2.3 Occlusion handling

Special care has to be taken in handling of the borders of moving objects. Here, an effect called occlusion is occurring, in which image parts are only visible in either of the two adjacent video frames. Although this effect is well known in literature, its solution is not trivial in practice. Earlier products incorporating MC-PRC did not apply special measures in occlusion areas, other than having a robust interpolation that reduces interpolation artifacts as much as possible.

Figure 7 Interpolation results using (left) motion compensated average, and (right) cascaded median.

This has been proven to be already highly valuable, as a simple interpolation may result in the highly visible artifacts shown in Figure 7 (left). Here, it is shown that, due to lack of occlusion handling, the interpolation frequently contains a mix of foreground and background object in occlusion areas, resulting in highly visible artifacts. A significant improvement can already be achieved by the aforementioned cascaded median, which reverts to non-motion compensated pixels in these areas.

Unfortunately, the drawbacks of this method become apparent if the background object contains details, as shown in Figure 8 (left). Now, the erroneous interpolation of the cascaded median produces a shadow-like artifact in the occlusion areas that, when observed in real-time, resembles a halo around moving objects.

To fundamentally solve the occlusion/halo artifacts, the algorithm has to perform a number of steps: 1) the occlusion areas have to be located, 2) the correct motion vectors have to be determined in these areas, and 3) the interpolation has to fetch data from only one of both video frames. The result of such an algorithm [7] is shown in Figure 8 (right), which illustrates the clear reduction of occlusion/halo artifacts.

Figure 8 Interpolation result using (left) cascaded median, and (right) extrapolation in occlusion regions.

2.4 Global fallback

Despite having met the above requirements, situations can still occur in which MC-PRC generates highly visible artifacts. This mainly holds for scenes that contain extremely irregular or erratic motion, or in which consecutive video frames show very little correspondences. The latter may be the case due to film edits (e.g. scene changes) or in synthetically generated content.

Figure 9 Interpolation results in the case of erratic motion without (left) and with (right) global fallback.

For such video content, picture-rate conversion using simple picture repetition is typically preferred over MC-PRC, as interpolation artifacts disappear, only to be replaced by judder. Perceptually, the human observer will have difficulty tracking the motion, making the re-appearance of judder a relatively minor issue, whilst interpolation artifacts tend to remain visible in such cases. The effects of this are illustrated in Figure 9.

In the decision on when to temporarily revert to picture repetition, multiple statistical cues may be used. For example, the noisiness of the motion vector field and/or the average residue can be used, as both may indicate the presence of large motion vector errors. Lastly, to prevent visible switching artifacts, the transition in and out of fallback may be performed gradually [8].

2.5 Large velocity range

Generally, larger object velocities result in a larger magnitude of the judder, and hence increase the relevance of MC. The implication of this is that the MC-PRC solution should be able to support sufficiently large velocity ranges. Although this does not pose significant algorithmic challenges, it does increase the architectural challenges as local buffer sizes have to be increased.

In most architectures, this impacts especially the vertical velocity range, as the video stream enters the system in a line-based fashion, and hence is also buffered in such a way. The consequence of an insufficiently large velocity range is shown in Figure 10, where break-up artifacts become visible (left).
of a 60Hz tickertape on top of film-content in 3:2 pull-down, as shown in Figure 11. Multiple options exist to handle this so-called hybrid content, such as globally processing the video’s dominant mode or not applying MC-PRC. Although neither option is necessarily superior, it is clear that special attention has to be paid to hybrid material in the design of the system.

3. Conclusions
MC-PRC plays a key role in improving the motion portrayal of LCD-TVs. Our learning curve started more than 20 years ago, leading to the first commercial consumer motion-compensated picture-rate conversion (MC-PRC) design in 1995. Since then, various designs have appeared on the consumer TV market. We like to emphasize that the sticker “MC-PRC” not automatically guarantees a high perceived picture quality, as a wide variety of requirements have to be met.

From an algorithmic point of view, special care has to be taken to achieve sufficient robustness, as MC-PRC has no unambiguous solution. This makes it an art more than a science to prevent visually annoying artifacts through local and global robustness measures.

Next to this, there are also architectural challenges, as e.g. the support of large velocities ranges which requires very large buffer sizes. Reducing these buffer sizes leads to an obvious cost-saving at a much decreased quality point. It still can have the sticker “MC-PRC” though…

Lastly, a far from trivial film mode detector has to ensure that the MC-PRC functionality can be applied to a wide variety of content. Again cost-savings are possible here, but it renders the functionality useless for many sources.

Only if all these requirements are met, and with our products we have shown this is feasible, our experience shows that the MC-PRC solution meets the picture quality requirements of both the mainstream and high-end TV segments.

4. References