ABSTRACT
There are now two generations of ICs for MC video format conversion (MC-VFC). Real-time DSP software for MC-VFC has recently been demonstrated, with the breakthroughs enabling this progress coming from motion estimation. The paper gives an overview.

INTRODUCTION
Motion estimation (ME) is a powerful way of improving a variety of video processing systems. Many methods have been developed for different applications over the last 30 years or so, including motion compensated (MC) filtering for noise reduction, MC prediction for coding, and MC interpolation for video format conversion (MC-VFC).

Predictive coding is the least demanding technique, and the first that led to widespread use of ME, as it is sufficient for coding that MC decreases the average prediction error. MC-VFC is probably the most demanding technique, as it requires estimation of true motion over a large range for picture rate conversion, and with a sub-pixel accuracy for de-interlacing.

Nevertheless, the last decade has brought us two generations of ICs for MC-VFC in consumer TV, and real-time DSP software for MC-VFC has also recently been demonstrated. All these concepts include ME with a quality that required substantial breakthroughs.

THE CRUCIAL DEVELOPMENTS
The first MEs required many calculations. Simplification was essential to allow use in consumer ICs. It is interesting that, although the operation count of the ME algorithms for MC-VFC has indeed decreased over time, their quality has improved. Intuitively, this corresponds to a trend seen in the techniques proposed, which range from pixel-based methods (pel-recursive, or gradient algorithms) [1], via block-based (the various flavours of block matching) [2], to object-based algorithms [3].

Block matching
This trend is in itself, however, not enough. There are fewer blocks than pixels in an image, but the huge number of candidates tested in a full search-block matcher (FS-BM) demands much more processing than the single calculation of an update in a pel-recursive ME, even if this calculation is repeatedly for every pixel. The search for simpler match criteria has converged via normalized cross correlation, then mean squared error, to the most popular summed absolute difference. This did not, however, reduce the operation count by the required orders of magnitude. The proposed alternatives with still lower hardware costs sacrificed too much performance.

A more significant reduction of the complexity resulted from efficient search techniques. For coding applications, this resulted in three-step search [4], logarithmic search [5], and one-at-a-time search techniques [6]. Unfortunately, all these methods decreased the already weak correlation with the true object motion.

Hierarchical methods [7,8] provided a breakthrough, as these not only reduced the operations count, but also improved the consistency of the vector field to a level at which it resembled the true motion. Derivatives of both hierarchical methods: the hierarchical block matching of [7] and the phase plane correlation (PPC) of [8] are found in professional MC-VFC equipment.

The breakthrough necessary for consumer ICs came from the introduction of the recursive search block matcher (RS-BM). Here, for objects larger than blocks and with inertia, the best candidate occurs in a spatio-temporal neighbourhood.

Moreover, a superior large-range true-motion ME can be realized with just three spatio-temporal predictors, a single random vector, and a well-chosen set of penalties added to the match criterion. It is this concept that is used in the 1st generation MC-VFC ICs [9,10].

In the 2nd generation [11], sub pixel accuracy was added [12] to enable the very demanding deinterlacing in addition to picture rate conversion. This generation further applied an extra candidate (next to the spatio-temporal prediction and the random candidate) derived from a global motion model which improves the performance in the event of camera movements [13].

1 The ICs are commercially available as type numbers SAA4990, SAA4991 and SAA4992.
Object based motion estimation

A natural extension of this parameter model capable of describing the (global) camera movements results from segmenting the image into individual objects and estimating motion parameters for each object. This follows the pixel-block-object trend, which is attractive because the number of blocks usually exceeds the number of objects by more than an order of magnitude. We can therefore hope for further potential reductions in complexity (Figure 1). Moreover, it guarantees the consistency of the vector field within objects.

The problem with object-based ME (OME) turns out to lie in the object segmentation. This is a far from trivial task that costs many more operations than the actual motion parameter estimation [14]. Here, the breakthrough came with the insight that it is possible to simply start a number of independent and different parameter estimators, and assign image portions to individual estimators. Although the assignment, which is based on a match criterion, occurs rather accidentally, the process converges towards a fair object segmentation provided that each individual parameter estimator is focused on the most suitable image sections. This is achieved by increasing the contribution to the match criterion for these sections, while decreasing the contribution from the rest of the picture. It is also interesting that the match optimization can be realized on some 1% of the luminance data and the assignment on a down-scaled image. This greatly reduces the operations count to a level where real-time DSP software becomes feasible.

SUMMARY

Progress in the field of ME has caused an evolution in algorithms along the path of pixel-, block-, and object-based methods. While operations count has decreased over time, quality has greatly increased, and the calculations have become more irregular. The hardware/software balance of the algorithms consequently moved to increased software content, while the most recent algorithms are implemented entirely in software running real time on a DSP.

REFERENCES