Nokia Exhibit 2007, p. 1
Amazon.com v. Nokia
IPR2024-01505

A Multi-Pass True Motion Estimation Scheme With Motion Vector Propagation for Frame Rate Up-Conversion Applications

Shen-Chuan Tai, Ying-Ru Chen, Zheng-Bin Huang, and Chuen-Ching Wang

***Abstract*—This paper presents a robust true motion estimation algorithm, designated as MPMVP (Multi-pass and Motion Vector Propagation), to enhance the accuracy of the motion vector fields in frame rate up-conversion applications. The MPMVP uses a multi-pass scheme to progressively refine approximate motion vectors to true motion vectors based upon the motion information acquired in previous passes. The multi-pass motion estimation process uses a large block size to detect the motion vectors within the objects themselves and small block sizes to detect the motion vectors along the object boundaries. Actually, the block size is progressively reduced during the search process. When the motion vector of a block is considered to be sufficiently accurate for motion estimation purposes, the block is said to be converged and the local motion vector search process terminates. A novel technique, referred to as motion vector propagation, is then applied to propagate the motion vector of the converged block to its neighboring blocks. This technique not only ensures that neighboring motion vectors within the same object have a high degree of spatial correlation, but also accelerates the convergence of the motion vectors in the neighboring blocks and therefore reduces the overall computational time and expense of the multi-pass motion vector search procedure. A novel distortion criterion is proposed to enhance the tolerance of the traditional sum-of-absolute-difference measurement technique applied in the motion estimation scheme to noise and shadow effects. The experimental results demonstrate that the proposed true motion estimation algorithm outperforms the traditional full search, 3DRS and TCSBP algorithms in terms of both the smoothness of the generated motion vector fields and the visual quality of the up-converted frames.**

***Index Terms*—Candidate motion vector, distortion criterion, frame rate up-conversion (FRUC), recursive search, true motion, variable block-size.**

I. Introduction

MODERN video formats have a wide variety of resolutions, aspect ratios, frame rates, and so forth. Typically, the frame rates of capturing devices vary between 24 and 60 Hz. However, the frame rates used by most display devices tend to range from 50 to 120 Hz, for example. Therefore, when playing back video sequences, it is necessary to interpolate some additional frames based on the original video sequence using some form of frame rate up-conversion (FRUC) algorithm. However, traditional schemes such as picture repetition, linear interpolation and other non-motion compensated methods generally produce visual artifacts such as trembling or blurring in the interpolated frames. The quality of the interpolated frames can be improved by combining the FRUC scheme with motion estimation (ME) techniques. Many complex algorithms for interpolating frames based upon the motion information of objects between successive frames have been proposed [1]–[15]. However, they tend to be highly time consuming and computationally expensive.

The ME schemes used in FRUC algorithms differ from the conventional ME techniques applied in coding standards. Generally speaking, ME algorithms used in frame rate up-conversion applications are termed "true ME" schemes in order to emphasize the importance of obtaining accurate motion vectors for all of the objects in the frame. True ME is employed not only in FRUC schemes, but also in such diverse applications as video encoding, [16], surveillance system tracking [17], de-interlacing schemes [18], error concealment in decoding systems [19], [20], and so forth. Briefly, the objective of video coding systems is to minimize the motion compensated residual. In block-matching ME algorithms, the features of moving objects are tracked using distortion criteria such as the sum of absolute difference (SAD) or the sum of square error (SSE). However, video sequences tend to contain high levels of noise, and, therefore, the motion vectors (MVs) extracted by minimal SAD/SSE-based coding systems do not always accurately reflect the true motions of the objects in the frame. In FRUC schemes, it is essential to obtain accurate motion vectors to ensure the quality of the interpolated frames, and therefore the ME algorithms used in such schemes require a robust ability to deal with the effects of noise. Unfortunately, the ME algorithms conventionally used in block-matching coding systems, e.g., full search, logarithmic search, diamond search, and so on, fail to address this requirement.

In frame rate up-conversion, the optimal image quality is obtained when the objects in the frames are interpolated along their respective motion trajectories. 3-D recursive search block matching (3DRS) [4] and temporal compensated motion estimation with simple block-based prediction (TCSBP) [2] have proposed ME algorithms based on predictive motion vectors (PMVs) to accelerate the block-matching speed and to enhance the degree of spatial correlation between neighboring blocks. In [3], the authors used an overlapped block-based motion estimation (OBME) scheme possessing enhanced ME capabil-

Manuscript received August 8, 2006; revised November 12, 2006, April 23, 2007, and October 20, 2007.

S.-C. Tai, Y.-R. Chen, and Z.-B. Huang are with the Department of Electrical Engineering, National Cheng Kung University, Tainan 701, Taiwan, R.O.C. (e-mail: cyr91d@dcmc.ee.ncku.edu.tw).

C.-C. Wang is with the Institute of Electrical Engineering, National Changhua University of Education, Changhua City 500, Taiwan, R.O.C.

Color versions of one or more figures are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JDT.2007.916014

1551-319X/$25.00 © 2008 IEEE
Authorized licensed use limited to: Emily Tannenbaum. Downloaded on August 11,2024 at 18:09:27 UTC from IEEE Xplore. Restrictions apply.

Nokia Exhibit 2007, p. 2
Amazon.com v. Nokia
IPR2024-01505

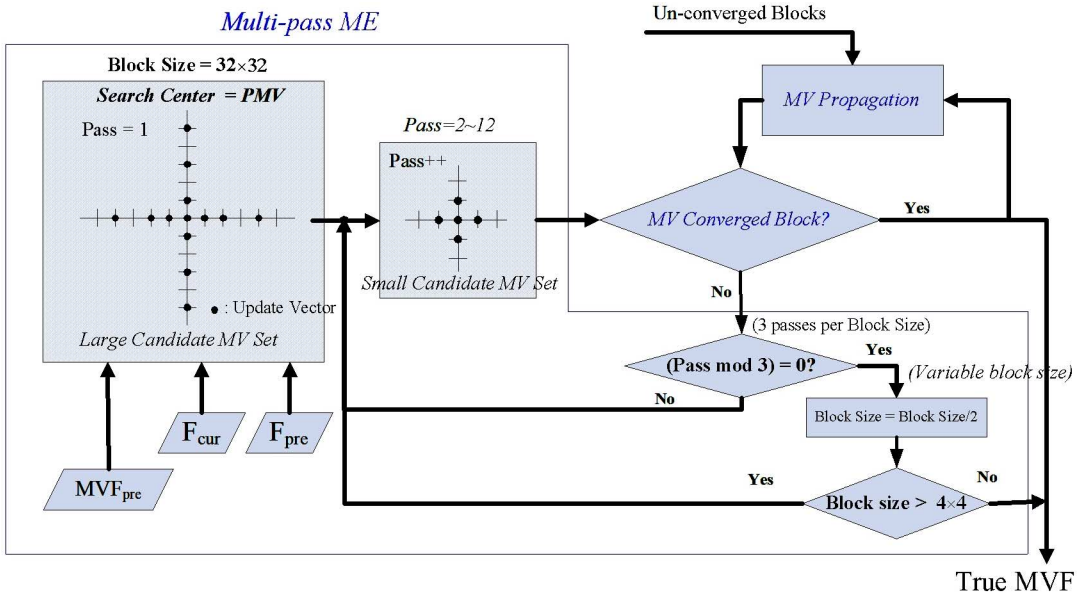


Fig. 1. Flow diagram of overall system.

ities and the ability to detect global motion through the use of a large-block matching mechanism. However, without accurate and sufficient information regarding the object shape, OBME schemes are unable to predict the motion of the object boundaries with any degree of accuracy, and therefore result in an incorrect motion aggregation in the motion vector field (MVF). A fundamental requirement of all true ME schemes is to achieve a high degree of spatial motion correlation inside a frame since all of the blocks within the same object must necessarily move in the same direction. Although several algorithms have been presented [4]–[6], [13] which employ some form of penalty function to enhance motion correlation, they are incapable of resolving the problem of incorrect motion propagation. In [10], [11], the MVs were established using hierarchical block-matching algorithms and were then propagated between hierarchical levels based on the regions of the objects. However, obtaining the regions of the objects requires a time-consuming segmentation process and the accuracy of the segmentation results cannot be guaranteed. In [14], a variable block size technique was used to enhance the ME accuracy. In this approach, the MVs were initially found for small blocks and were then used as the criteria for merging neighboring blocks. Having merged neighboring blocks, the MV search procedure was then repeated. However, this method is not only computationally expensive, but also achieves only limited success since small blocks cannot provide sufficient information regarding the global characteristics of an object and thus the reliability of the merging process cannot be ensured.

Accordingly, this paper presents a robust true ME scheme for FRUC applications. The remainder of this paper is organized as follows. Section II presents an overview of the major components of the proposed scheme, including the multiple-pass, the MV convergence decision technique, the MV convergence propagation scheme, and the distortion criterion used to suppress the effects of noise and shadow in the interpolated frames. Section III presents the experimental results obtained by the proposed true ME scheme for the MVFs of several well-known video test sequences, and compares the performance of the proposed scheme with that of the full search, 3DRS and TCSBP schemes when applied to frame rate up-conversion applications. Finally, Section IV presents some brief conclusions.

II. Multi-Pass Motion Estimation Strategy and MV Propagation Rule for True Motion Estimation

The proposed ME scheme searches for the true MVs using a multi-pass strategy in which the block size is progressively reduced from 32×32 to 4×4. After each pass, a convergence decision rule is applied to determine whether or not the MV identified for the block is sufficiently accurate. If the MV has acceptable accuracy, the block is said to be converged and the local MV search procedure is terminated. As indicated above, the largest block size used in the multi-pass ME scheme is 32×32. When an object is bigger than this block size, a technique referred to as MV propagation can be applied to spread the MV of the converged block to all of the blocks within the object. Fig. 1 presents an overall view of the multi-pass MV search procedure and associated MV propagation routine.

A. Multi-Pass Motion Estimation Strategy

The basic principle of the multi-pass ME strategy proposed in this study is to utilize the motion information relating to neighboring blocks generated in the previous pass to gradually refine the accuracy of the MV in the current block. The multi-pass search scheme executes a maximum of 12 passes. In the first three passes, the MV search is performed using a block size of 32×32. However, in the next three passes, the block size is reduced to 16×16, while in the final two sets of three passes, the block size is further reduced to 8×8 and 4 × 4, respectively. The MV for block $\overrightarrow{X}$ found in frame t in the ith pass is defined as $\overrightarrow{\mathrm{MV}}^i(\overrightarrow{X}, t)$. The MV search in block $\overrightarrow{X}$ in the i th pass is performed in accordance with (1). As shown, the search is confined to the vectors $\overrightarrow{V}$ defined in the candidate vector set $CS^i(\overrightarrow{X}, t)$, i.e., a full search is not performed. As discussed later, the term

Authorized licensed use limited to: Emily Tannenbaum. Downloaded on August 11,2024 at 18:09:27 UTC from IEEE Xplore. Restrictions apply.

Nokia Exhibit 2007, p. 3
Amazon.com v. Nokia
IPR2024-01505

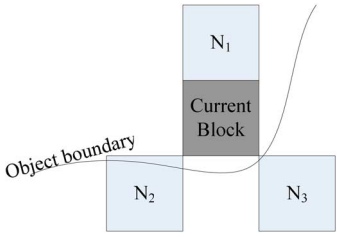


Fig. 2. Geometric relation between motion edge and block position.

$e(\vec{V},\vec{X},t)$ in (1) is a new distortion criterion based upon the conventional SAD distortion criterion.

$$\overrightarrow{MV}^{i}(\vec{X},t) = \arg \min_{\vec{V}\in CS^{i}(\vec{X},t)} (e(\vec{V},\vec{X},t)), \qquad 1 \le i \le 12. \tag{1}$$

The candidate vector set $CS^i(\vec{X},t)$ is comprised by a single PMV and a set of update vectors. Let $\vec{X}$ represent the block location and let t be the frame number. The candidate vector set for this particular block in the ith pass are then defined as

$$CS^i(\vec{X},t) = \left\{ \vec{V} \,|\, \vec{V} = \overrightarrow{\text{PMV}}^{i}(\vec{X}) + \vec{\text{U}}, \vec{\text{U}} \in US^i \right\} \tag{2}$$

where $\overrightarrow{\text{PMV}}^{i}(\vec{X})$ is the PMV of block $\vec{X}$ and is chosen from the MVs of the neighboring blocks and US^i is a set of update vectors. Every motion vector $\vec{V}$ in the candidate vector set is comprised by a single PMV and a update vector $\vec{U}$ selected from update vector set US^i.

1) PMV: Although the MVs of all eight neighboring blocks can theoretically be used as PMV predictors for the current block, evaluating each of these predictors in order to establish the most suitable PMV among them is computationally expensive and time consuming. Therefore, this study uses the geometric relationship proposed in [2] describing the correlation between the current block position and the motion edge to reduce the number of PMV predictors to be evaluated. As shown in Fig. 2, at least one of the three neighboring blocks and the current block must belong to the same object no matter where the object boundary crosses, and, therefore, the number of PMV predictors which must be considered can be reduced from 8 to just 3. The scheme presented in [2] involved a single ME pass, and therefore employed a single PMV predictor pattern. However, as described above, the current search scheme executes a maximum of 12 passes. In general, different PMV predictor patterns may result in different levels of accuracy. Therefore, the current multi-pass ME scheme considers 12 different PMV predictor patterns, as shown in Fig. 3.

Each pass in the proposed ME scheme is performed using one of the PMV predictor patterns (see Fig. 3) and then one MV in the selected PMV pattern which has minimal distortion is selected as the final PMV of this pass. In traditional single-pass ME algorithms, the current block has no information regarding the motion of the succeeding blocks because their MVs are not obtained yet. However, in multi-pass ME algorithms, the motion information of all blocks in current pass can be referred in future passes.

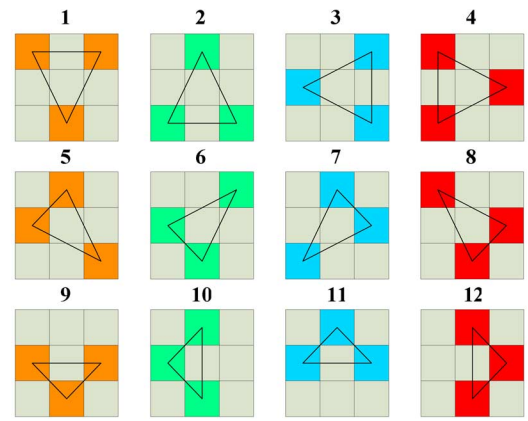


Fig. 3. PMV predictor patterns.

2) Update Vector Set: In (2), the term US^i represents the set of update vectors in the ith pass and has the form

$$US^i = \begin{cases} US_{\text{large}}, & i = 1 \\ US_{\text{small}}, & 2 \le i \le 12 \end{cases}$$
$$US_{\text{small}} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \pm 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \pm 1 \end{pmatrix} \right\}$$
$$US_{\text{large}} = US_{\text{small}} \cup \left\{ \begin{pmatrix} \pm 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \pm 3 \end{pmatrix}, \begin{pmatrix} \pm 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \pm 5 \end{pmatrix} \right\}. \tag{3}$$

The update vector sets employed in the proposed algorithm vary between different passes. The proposed scheme uses a large update vector set (i.e., a greater number of larger MVs) in the initial passes when employing a large block size to detect coarse motions in the frame. When the update vectors have a large range, e.g., ±(5,5), finding good approximations to the true MVs in blocks which cross object boundaries is straightforward. Furthermore, in the first pass, the motion information of neighboring blocks is simply copied from the previous frame and therefore there is no guarantee that this information is reliable. As a result, the PMV obtained for the block is not necessarily an accurate representation of the true MV. However, using a large update vector set increases the probability of obtaining a good approximation to the true MV. In later passes of the search procedure, a highly precise PMV with a small update vector set (i.e., fewer and smaller vectors) is used to refine these initial non-true motion estimates into true motion estimates.

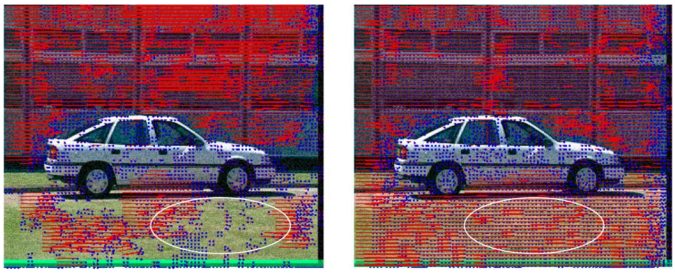
As indicated earlier, the term $e(\vec{V},\vec{X},t)$ in (1) is a new distortion criterion based on the conventional SAD distortion criterion and has the form

$$e(\vec{V},\vec{X},t) = \sum_{i,j\in \text{block}} \text{weightAD}(\vec{V},\overrightarrow{X_{ij}},t)/\text{blocksize}$$
$$\text{AD}(\vec{V},\overrightarrow{X_{ij}},t) = \text{abs}(I(\overrightarrow{X_{ij}},t) - I(\overrightarrow{X_{ij}} + \vec{V}, t-T))$$
$$\text{weight} = \begin{cases} 0, & \text{AD}(\vec{V},\overrightarrow{X_{\text{ij}}},t) < 5 \\ 1, & 5 \le \text{AD}(\vec{V},\overrightarrow{X_{\text{ij}}},t) < 30 \\ 2, & \text{otherwise} \end{cases} \tag{4}$$

where $I(\vec{X}_{\text{ij}},t)$ is the gray level of a pixel located at coordinates (i,j) in block $\vec{X}$ of frame t and abs indicates absolute value. According to Weber's law, human eyes have a noise threshold

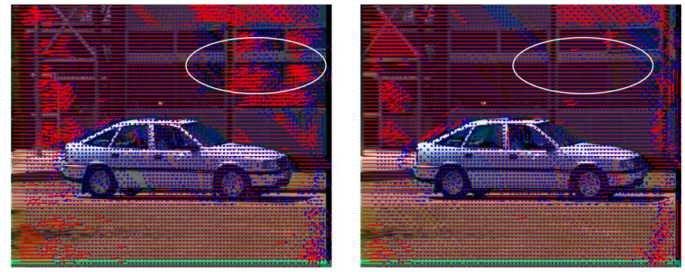
Authorized licensed use limited to: Emily Tannenbaum. Downloaded on August 11,2024 at 18:09:27 UTC from IEEE Xplore. Restrictions apply.

Nokia Exhibit 2007, p. 4
Amazon.com v. Nokia
IPR2024-01505



(a) (b)

Fig. 4. MVFs of frames with Gaussian noise: (a) MVF produced by proposed algorithm with traditional SAD distortion criterion, and (b) MVF produced by proposed algorithm with new distortion criterion.



(a) (b)

Fig. 5. Up-converted frames from VECTRA sequence. (a) MVF produced by 3DRS with penalty function. (b) MVF produced by 3DRS with proposed distortion criterion.

of approximately 5 grayscale value (for contrast) in pictures. For experience, an edge often leads to a gradient greater than 30 grayscale value. Therefore, if the absolute difference (AD) between current pixel and referred pixel in previous frame is less than 5, the human eye will ignore the distortion. However, if it is greater than 30, the distortion will be very noticeable. Hence, as shown in (4), the distortion criterion applied in the current ME search scheme has the form of a piecewise linear function. In general, noise has a particularly serious effect on the accuracy of the MV search process in smooth areas of the frame, i.e., background area. Therefore, the piecewise linear function in (4) allows small distortions $(\mathrm{AD} < 5)$ caused by noise to be ignored while simultaneously placing more emphasis on the large distortions $(\mathrm{AD} \geq 30)$ caused by the edges, and therefore the accuracy of the MV is enhanced.

Fig. 4 shows the MVFs obtained by the proposed ME scheme using the conventional SAD distortion criterion and the new distortion criterion given in (4), respectively, for a frame with Gaussian noise. Comparing the two frames, it is apparent that the improved noise suppression characteristics of the proposed distortion criterion yield more accurate MVs vectors in the smooth regions of the image, i.e., on the grass. Fig. 5 presents two up-converted frames from the VECTRA test sequence obtained by the 3DRS search scheme using the original penalty function and the new distortion criterion, respectively. It can be seen that the MVF produced by 3DRS using the new distortion criterion is more smooth. For example, in Fig. 5(b) more motion vectors of the window area indicate to the same direction than in Fig. 5(a). In other words, the proposed distortion criterion is applicable not only to the multi-pass scheme presented in the current study, but also to existing true ME algorithms.

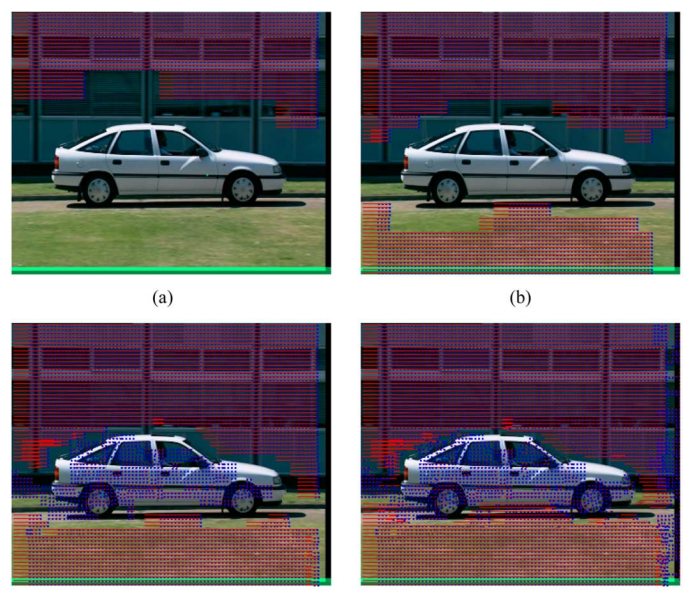
B. Decision Rules for MV Converged Block

To reduce the computational complexity and expense of the proposed ME search scheme, processing within a local block terminates as soon as the block is converged, irrespective of the number of passes which have been completed. Furthermore, the degree of motion correlation among neighboring blocks is enhanced by using a MV propagation technique that will be described in Section II-C. The degree of spatial motion correlation between neighboring blocks is determined in accordance with

$$\Delta\mathrm{MV}(\overrightarrow{C},\overrightarrow{N}) = \mathrm{abs}(\mathrm{MV}(\overrightarrow{C},t).x - \mathrm{MV}(\overrightarrow{N},t).x) + \mathrm{abs}(\mathrm{MV}(\overrightarrow{C},t).y - \mathrm{MV}(\overrightarrow{N},t).y) \quad (5)$$

(a) (b)

Fig. 6. (a) Proposed ME scheme without convergence propagation and (b) proposed ME scheme with convergence propagation.



(a) (b) (c) (d)

Fig. 7. Motion vectors of converged Blocks (MV is shown by block-size 4×4). (a) 3rd Pass: block size is 32 × 32, (b) 6th Pass: block size is 16×16. (c) 9th Pass: block size is 8×8. (d) 12th Pass: block size is 4×4.

where $\overrightarrow{C}$ and $\overrightarrow{N}$ are the current block and a neighboring block in spatial motion correlation test range, respectively, and x and y are the horizontal and vertical MV components. True motion is obtained when the distortion value is low, i.e., $e(\overrightarrow{C},\overrightarrow{X},t) < T_1$ and the degree of motion correlation between the current block and all of its adjacent blocks in the spatial motion correlation

Authorized licensed use limited to: Emily Tannenbaum. Downloaded on August 11,2024 at 18:09:27 UTC from IEEE Xplore. Restrictions apply.

Nokia Exhibit 2007, p. 5
Amazon.com v. Nokia
IPR2024-01505

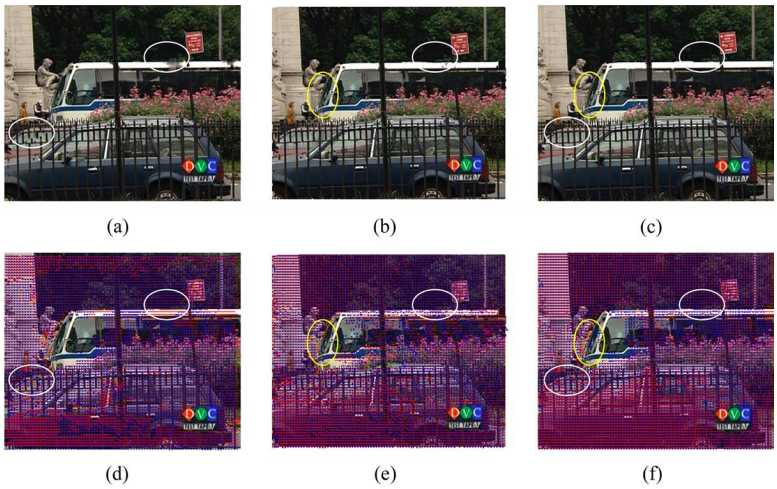


Fig. 8. (a), (d) Interpolated frame from BUS sequence and corresponding MVF obtained from 3DRS, (b), (e) Interpolated frame from BUS sequence and corresponding MVF obtained from TCSBP. (c), (f) Interpolated frame from BUS sequence and corresponding MVF obtained from proposed MPMVP algorithm.

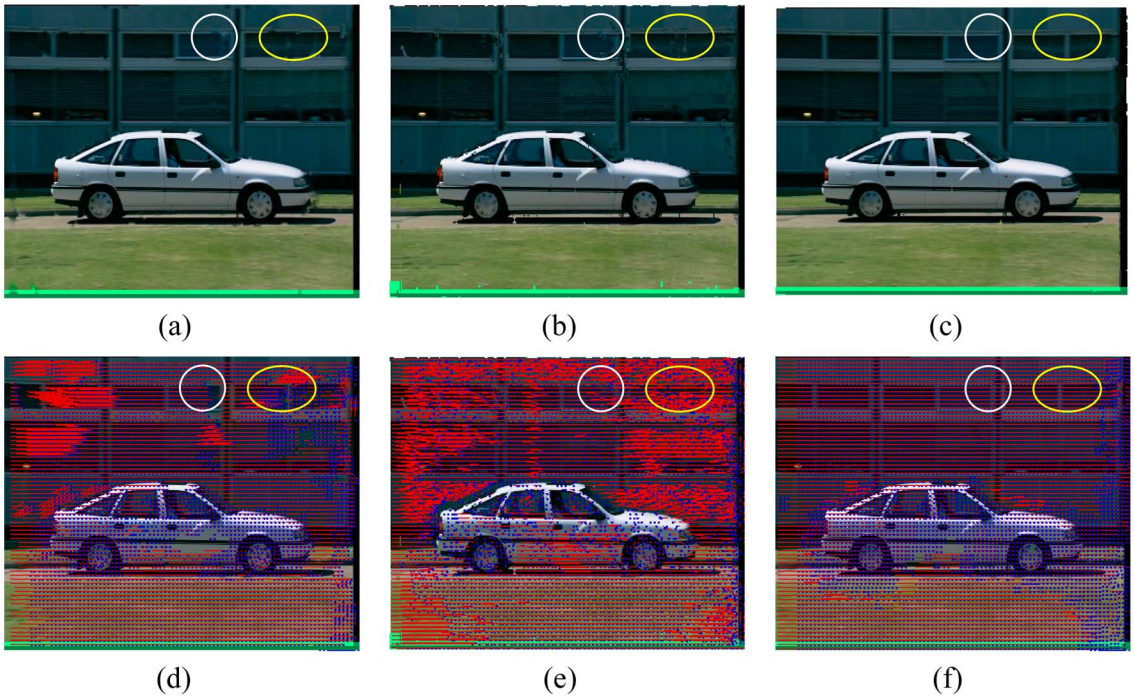


Fig. 9. (a), (d) Interpolated frame from VECTRA sequence and corresponding MVF obtained from 3DRS. (b), (e) Interpolated frame from VECTRA sequence and corresponding MVF obtained from TCSBP. (c), (f) Interpolated frame from VECTRA sequence and corresponding MVF obtained from proposed MPMVP algorithm.

test range is high, i.e., $\Delta\mathrm{MV}(\overrightarrow{C}, \overrightarrow{N}) < T_2$, where T_1 is a distortion threshold and T_2 is a threshold value designed to enable sufficient activity for the MV search to prevent the clustering of inappropriate MVs. In other words, the current scheme considers both the distortion value and the degree of motion correlation when searching for true MV. This approach differs from existing ME algorithms, which typically terminate the MV evaluation process as soon as an acceptable distortion value has been obtained.

The threshold for the distortion value T_1 in ith pass is set as $5+i/3$ in accordance with experience. Furthermore, the spatial motion correlation test ranges used in the current algorithm are 3×3 and 5×5, respectively. Since there are three passes during specific-block-size ME, the larger test range is applied in the first two passes, while the smaller test range is used in the final pass. Meanwhile, T_2 is assigned a value of 0 for block sizes larger than 8×8 and a value of 1 for smaller blocks. Note that this threshold is deliberately assigned a stricter value for larger blocks to prevent them from converging prematurely.

C. MV Propagation Method

The MV propagation technique employed in the current scheme is designed to enhance the degree of motion correlation in large textureless regions of the frame. When a block converges, i.e., it has a high degree of spatial motion correlation and a low distortion, its MV is automatically offered to the neighboring unconverged blocks as a potential MV for their own motions. If the difference between the original distortion value of the neighboring block and the new distortion value of the block produced by accepting the propagated MV is

Authorized licensed use limited to: Emily Tannenbaum. Downloaded on August 11,2024 at 18:09:27 UTC from IEEE Xplore. Restrictions apply.

Nokia Exhibit 2007, p. 6
Amazon.com v. Nokia
IPR2024-01505

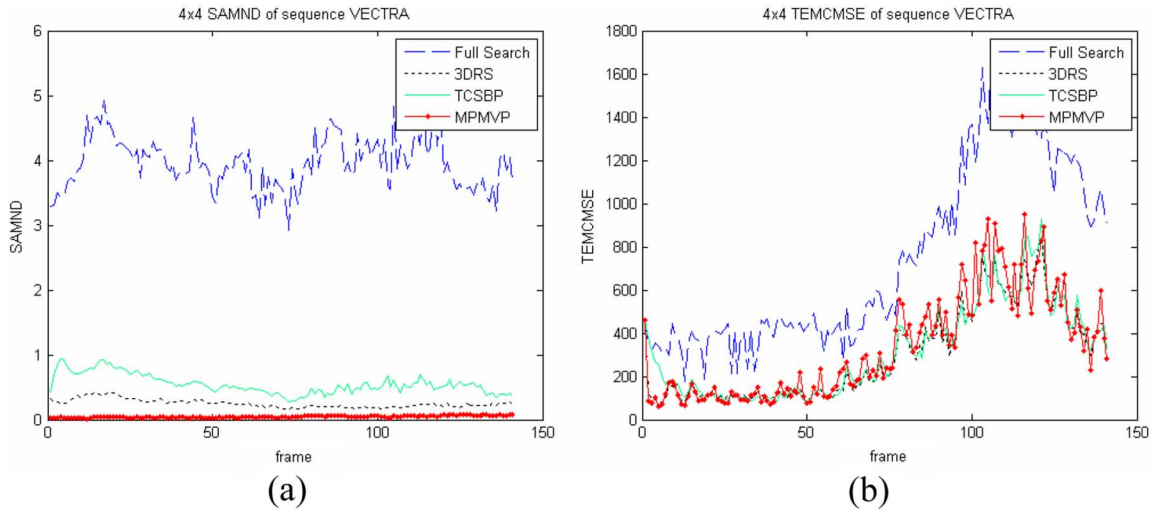


Fig. 10. SAMND and TEMC-MSE comparisons for full search (± 32), 3DRS, TCSBP and proposed algorithm. (Note results relate to VECTRA test sequence).

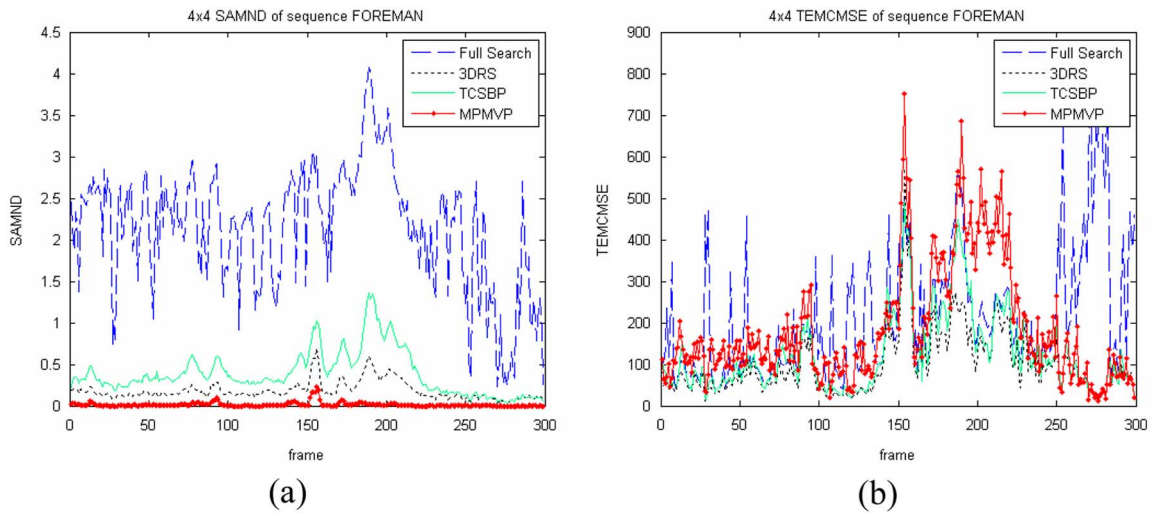


Fig. 11. SAMND and TEMC-MSE comparisons for full search (± 16), 3DRS, TCSBP and proposed algorithm. (Note results relate to FOREMAN test sequence).

sufficiently small, i.e., $15 - 2/3^{*}i$, the neighboring block accepts the propagated MV as its own MV. Having accepted the propagated MV, the block then checks to see whether or not the MV causes it to become converged. When a block has converged, irrespective of whether or not convergence is induced by the propagated MV, the same MV is propagated to the neighbors of the new converged block.

The convergence propagation technique is used to expand size of the region, which has the similar MV, into the size of object. Convergence propagation ensures that the MVs of the background regions of a frame have a high degree of spatial correlation. In other words, the background appears as a single, continuous region with no obvious areas of demarcation, even across strong edge areas, as shown in Fig. 6.

III. Experimental Results

In the figures discussed in the following, the MVFs and interpolated frames are based on a block size of 4×4. For example, a MV of block size 32×32 will be presented by 64 MVs of block size 4×4. Each MV is colored in red, originates from the center of the 4×4 block and terminates with a solid blue point. Fig. 7 presents the MVs of converged blocks obtained by the proposed ME scheme with different block sizes. The MVs of the background regions of the frame are found in the earlier passes of the ME search process performed using a larger block size. The use of a variable block size in the search procedure enables the block size of MV to gradually fit the contour of the objects in each pass and therefore the accuracy of the identified MVs is greatly enhanced. As a result, Fig. 7 demonstrates that the algorithm successfully divides the frame into two distinct objects, namely the background and the moving car.

In general, interpolated (i.e., up-converted) frames are generated using a motion compensation technique. Some uncovered regions (i.e., regions with no MVs) can be interpolated based on neighboring available pixels. Since the areas of these uncovered regions usually are small, we use spatial bilinear interpolation to interpolate uncovered regions in motion compensation step of this paper. Figs. 8 and 9 present the MVFs and corresponding interpolated frames obtained using the 3DRS [4], TCSBP [2] and proposed MPMVP ME algorithms for the BUS and VECTRA video sequences, respectively. Note that in every case, the test sequences are in a CIF format, have an original frame rate of 30 frames/s, and are up-converted to a rate of 60 frames/s. Observing these figures, it is apparent that the 3DRS and TCSBP

Authorized licensed use limited to: Emily Tannenbaum. Downloaded on August 11,2024 at 18:09:27 UTC from IEEE Xplore. Restrictions apply.

Nokia Exhibit 2007, p. 7
Amazon.com v. Nokia
IPR2024-01505

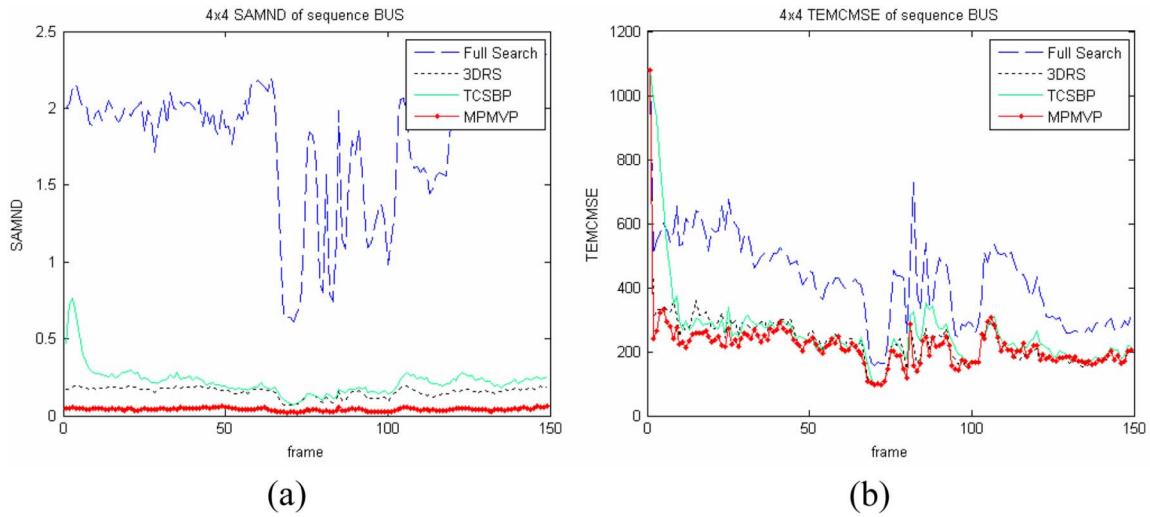


Fig. 12. SAMND and TEMC-MSE comparisons for full search (± 16), 3DRS, TCSBP and proposed algorithm. (Note results relate to BUS sequence).

schemes generate incorrect MVs in the smooth, expansive background regions of the frame because they fail to propagate the MVs efficiently among neighboring blocks.

The quality of the interpolated frames produced in FRUC applications is generally assessed using subjective visual quality measure. However, in [2], the authors proposed two new evaluation criteria with which to measure the motion correlation of interpolated frames. The first criterion was the spatial consistency indicator (i.e., the sum of absolute minimum neighboring difference, designated as SAMND), while the second criterion related to the temporal continuity of the motion field between successive frames (i.e., temporal extending motion compensation mean square error, designated as TEMC-MSE). SAMND evaluates the MV differences between neighboring blocks in accordance with

$$\mathrm{SAMND} = \sum_{\overrightarrow{X} \in \mathrm{MVF}_t} \min(\mathrm{abs}(\mathrm{Vec}_{\vec{X},t} - \mathrm{Vec}_{\vec{X}+\vec{\delta},t})) \quad (6)$$

where $\overrightarrow{X}$ is the block location, *Vec* is the MV of the block, and δ is the 3×3 neighboring region. Meanwhile, the TEMC-MSE criterion measures the temporal MV ductility and focuses on the smooth regions of the frame such as the background where the MVs in the previous frame are also valid for the current frame. The TEMC-MSE criterion has the form

$$\mathrm{TEMC-MSE} = \frac{\sum_{\vec{P}} (I(\overrightarrow{P}, t) - I(\overrightarrow{P} - \mathrm{Vec}_{\vec{P},t-T}, t-T))^2}{M \times N} \quad (7)$$

where

I	luminance component of a frame;
$\overrightarrow{P}$	pixel position;
$M \times N$	total number of sampling pixels in the frame;
$\mathrm{Vec}_{\vec{P},t-T}$	MV of position $\overrightarrow{P}$ at frame $t-T$, and can be estimated from the luminance fields $I(\overrightarrow{P}, t-T)$ and $I(\overrightarrow{P}, t-2T)$.

Figs. 10–12 compares the SAMND and TEMC-MSE metrics for the full search, 3DRS, TCSBP and proposed MPMVP for the VECTRA sequence (characterized by fast and complicated motions), FOREMAN sequence (characterized by low motions) and BUS sequences (characterized by complicated motions), respectively. In general, Figs. 10–12 show that the MVs generated by the proposed algorithm have the highest degree of spatial correlation of all the MVs produced by the four algorithms. However, in Fig. 11(b), the proposed algorithm has the poorest TEMC-MSE performance. The poor performance arises because the proposed distortion criterion deliberately ignores small distortion values in order to enhance the noise resistance. However, when checking the up-converted FOREMAN sequence by human eyes, it is found that the proposed algorithm generates the highest quality interpolated frames of the four algorithms. For example, Fig. 13 shows frames 154 and 155 of the original FOREMAN sequence together with the corresponding interpolated frames generated using the full search, 3DRS, TSCBP and proposed MPMVP schemes, respectively. It can be seen the visual quality of MPMVP is the best. [Note that these particular frames are deliberately chosen here because frame 154 has a high TEMC-MSE value in Fig. 11(b)]. These figures also show that the full search scheme has the poorest performance of the four algorithms in terms of both the SAMND and the TEMC-MSE metrics since it is not specifically designed to focus on motion correlation.

Although the ME scheme proposed in this study is a recursive motion search method, the convergence decision rules and MV propagation technique greatly reduce its computational time and complexity. For full search, the computation is decided by search range, i.e., for search range n, the computation is $(2n+1)^2$ ADs/pixel, where n is often set as 7, 15, 31, or even 63. For a worst-case in MPMVP, the computation is 101 ADs/pixel for MPMVP. Because in the first pass, the proposed algorithm takes the motion vector in the previous field as the PMV and then evaluates the 13 search points defined in $\mathrm{US}_{\mathrm{large}}$ [see (3)]. In the following 11 passes, the algorithm evaluates 3 PMV predictors and each of the 5 search points defined in $\mathrm{US}_{\mathrm{small}}$. Therefore, the maximal computational complexity is $13+(11^*(3+5)) = 101$ ADs/pixel. Bus in general cases, the actually evaluated search points are much less than 101 ADs/pixel.

Authorized licensed use limited to: Emily Tannenbaum. Downloaded on August 11,2024 at 18:09:27 UTC from IEEE Xplore. Restrictions apply.

Nokia Exhibit 2007, p. 8
Amazon.com v. Nokia
IPR2024-01505

Fig. 13. (a), (b) Frames 154 and 155 in FOREMAN sequence. (c)-(f) Frames interpolated by full search, 3DRS, TCSBP and proposed algorithm, respectively.

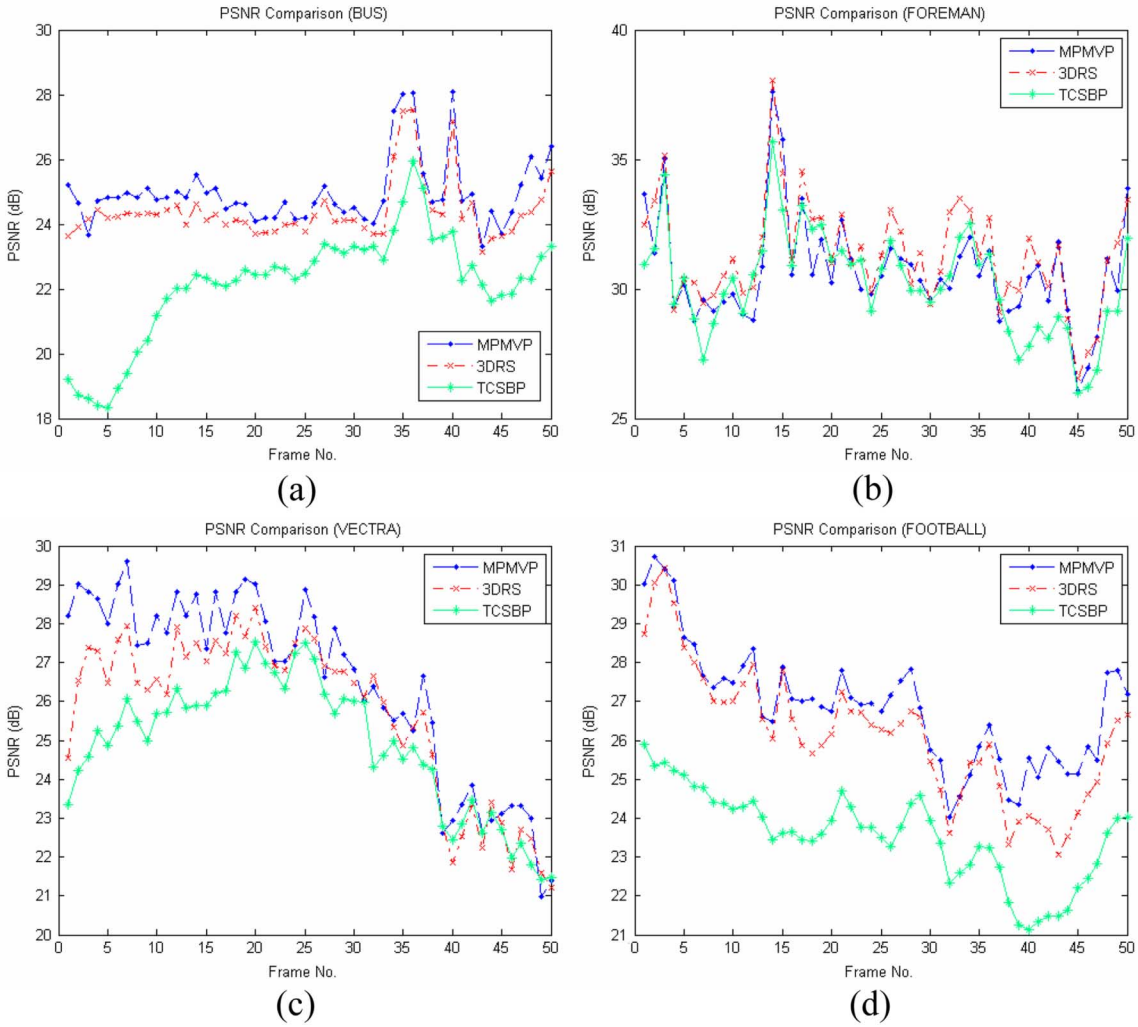


Fig. 14. PSNR comparisons on up-converted test sequences. (a) BUS. (b) FOREMAN. (c) VECTRA. (d) FOOTBALL. (a), (b), (c) are in CIF format and (d) is in de-interlaced D1 format.

Authorized licensed use limited to: Emily Tannenbaum. Downloaded on August 11,2024 at 18:09:27 UTC from IEEE Xplore. Restrictions apply.

Nokia Exhibit 2007, p. 9
Amazon.com v. Nokia
IPR2024-01505

TABLE I
MOTION ESTIMATION TIMES SPENT IN 3DRS, TCSBP AND PROPOSED MPMVP ALGORITHM

Time (sec/frame) / Sequence	3DRS	TCSBP	MPMVP
BUS (CIF)	0.317	0.037	0.145
FOREMAN (CIF)	0.315	0.036	0.074
VECTRA (CIF)	0.305	0.038	0.116
FOOTBALL (D1)	2.186	0.264	1.143

We have tested the motion search time and PSNR performance for 3DRS, TCSBP and the proposed MPMVP scheme for three CIF format test sequences. The frame rate of these sequences are firstly down-sampled from 30 to 15 fps and then up-converted to 30 fps. We calculate the PSNR between the up-converted frames and the original frames in 30 fps sequence. The PSNR curves of the first 50 up-converted frames are shown in Fig. 14, while the corresponding motion estimation times are listed in Table I. In order to test the performance of the proposed MPMVP in larger resolution sequences, one de-interlaced D1 sequence (50 fps), FOOTBALL, is down-sampled from 50 to 25 fps and then up-converted to 50 fps. The experimental result of FOOTBALL is also shown in Fig. 14 and Table I. From those data, it can be seen that using our MPMVP scheme in frame rate up-conversion applications not only can achieve higher quality up-converted frames, but also can save significant computation.

IV. CONCLUSION

This paper has presented a new highly-accurate, low-complexity ME scheme for FRUC applications. The proposed algorithm is based upon a multi-pass, variable block-size approach and features a new noise-resistant distortion criterion to improve the quality of the interpolated frames. Furthermore, efficient MV convergence criteria and MV propagation techniques are employed to reduce the overall computational time and cost of the scheme. The experimental results have shown that the proposed ME scheme not only achieves a higher MV accuracy than existing ME schemes, but also yields significant computational savings. The quality of our true MV generated up-converted frames are better than other ME algorithm generated frames.

REFERENCES

[1] B.-T. Choi, S.-H. Lee, and S.-J. Ko, "New frame rate up-conversion using bi-directional motion estimation," *IEEE Trans. Consumer Electron.*, vol. 46, no. 3, pp. 603–609, Aug. 2000.
[2] J. Wang, D. Wang, and W. Zhang, "Temporal compensated motion estimation with simple block-based prediction," *IEEE Trans. Broadcast.*, vol. 49, no. 3, pp. 241–248, Sep. 2003.
[3] H. Taehyeun, L. Seongjoo, and K. Jaeseok, "Motion compensated frame interpolation by new block-based motion estimation algorithm," *IEEE Trans. Consumer Electron.*, vol. 50, no. 2, pp. 752–759, May 2004.
[4] G. de Haan, P. W. A. C. Biezen, H. Huijgen, and O. A. Ojo, "True-motion estimation with 3-D recursive search block matching," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 5, pp. 368–379, Oct. 1993.
[5] R. Braspenning and G. de Haan, "True-motion estimation using feature correspondence," *SPIE, Proc. VCIP*, pp. 396–407, Jan. 2004.
[6] G. de Haan and P. W. A. C. Biezen, "An efficient true-motion estimator using candidate vectors from a parametric motion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 1, pp. 85–91, Feb. 1998.
[7] A. Pelagotti and G. de Haan, "A new algorithm for high quality video format conversion," in *2001 Proc. Int. Conf. on Image Process.*, Oct. 2001, vol. 2, pp. 375–378.
[8] S.-H. Lee, Y.-C. Shin, Y. Seungjoon, H.-H. Moon, and R.-H. Park, "Adaptive motion-compensated interpolation for frame rate up-conversion," *IEEE Trans. Consumer Electron.*, vol. 48, no. 3, pp. 444–450, Aug. 2002.
[9] S.-H. Lee and S. Yang, "Adaptive motion-compensated frame rate up-conversion," *Electron. Lett.*, vol. 38, no. 10, pp. 451–452, May 2002.
[10] B.-W. Jeon, G.-I. Lee, S.-H. Lee, and R.-H. Park, "Coarse-to-fine frame interpolation for frame rate up-conversion using pyramid structure," *IEEE Trans. Consumer Electron.*, vol. 49, no. 3, pp. 499–508, Aug. 2003.
[11] D. Wang and D. Lauzon, "Hybrid algorithm for estimating true motion fields," *Opt. Eng.*, vol. 39, no. 11, pp. 2876–2881, Nov. 2000.
[12] K. Hilman, H.-W. Park, and Y.-M. Kim, "Using motion compensated frame-rate conversion for the correction of 3:2 pulldown artifacts in video sequences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 6, pp. 869–877, Sep. 2000.
[13] P. Ishwar and P. Moulin, "On spatial adaptation of motion-field smoothness in video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 6, pp. 980–989, Sep. 2000.
[14] S. Fujiwara and A. Taguchi, "Motion-compensated frame rate up-conversion based on block matching algorithm with multi-size blocks," in *Proc. 2005 Int. Symp. on Intell. Signal Process. Commun. Syst.*, Dec. 2005, vol. 13–16, pp. 353–356.
[15] B.-D. Choi, J.-W. Han, C.-S. Kim, and S.-J. Ko, "Frame rate up-conversion using perspective transform," *IEEE Trans. Consumer Electron.*, vol. 52, no. 3, Aug. 2006.
[16] O. EePing, H. Wang, and P. Xue, "Video coding based on true motion estimation," in *Proc. 2003 IEEE Int. Conf. on Acoust., Speech, and Signal Process.,*, Apr. 2003, vol. 3, no. 6–10, pp. 409–12.
[17] J. E. Santos Conde, A. Teuner, and B. J. Hosticka, "Hierarchical locally adaptive multigrid motion estimation for surveillance applications," in *Proc. 1999 IEEE Int. Conf. on Acoust., Speech, Signal Process.*, Mar. 1999, vol. 6, pp. 3365–3368.
[18] G. de Haan, "IC for motion-compensated de-interlacing, noise reduction, and picture-rate conversion," *IEEE Trans. Consumer Electron.*, vol. 45, no. 3, pp. 617–624, Aug. 1999.
[19] J.-W. Suh and Y.-S. Ho, "Error concealment technique based on optical flow," *Electron. Lett.*, vol. 38, no. 18, pp. 1020–1021, Aug. 2002.
[20] J.-Y. Pyun, J.-S. Lee, J.-W. Jeong, J.-H. Jeong, and S.-J. Ko, "Robust error concealment for visual communications in burst-packet-loss networks," *IEEE Trans. Consumer Electron.*, vol. 49, no. 4, pp. 1013–1019, Nov. 2003.

Shen-Chuan Tai received the B.S. and M.S. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1982 and 1986, respectively, and received the Ph.D. degree in computer science from National Tsing Hua University, Hsinchu, Taiwan, R.O.C., in 1989.

He is currently a professor in the Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan. His teaching and research interests include data compression, DSP, VLSI array processors, computerized electrocardiogram processing, and multimedia systems.

Authorized licensed use limited to: Emily Tannenbaum. Downloaded on August 11,2024 at 18:09:27 UTC from IEEE Xplore. Restrictions apply.

Nokia Exhibit 2007, p. 10
Amazon.com v. Nokia
IPR2024-01505

Ying-Ru Chen received the B.S. degree in education from the Department of Information and Computer Education, National Taiwan Normal University, Taipei, Taiwan, R.O.C., in 2000, and the M.S. degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, R.O.C., in 2002, and where she is currently working toward the Ph.D. degree candidate in the Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C.

Her research interests include video signal processing, data compression, multimedia systems, and digital watermarking.

Zheng-Bin Huang received the B.S. degree in electrical engineering and the B.S. degree in mechanical engineering from National Cheng Kung University, Tainan, Taiwan, R.O.C., both in 2004, and the M.S. degree in electrical engineering from the National Cheng Kung University, Tainan, Taiwan, R.O.C., in 2006.

His research interests include video compression, image processing, and fast true motion estimation algorithms.

Chuen-Ching Wang received B.S., M.S., and Ph.D. degrees in electrical engineering from the National Cheng Kung University, Taiwan, R.O.C., in 1986, 1994, and 2003, respectively.

He was a section chief at Directorate General Telecommunications in the Ministry of Transportation and Communications. He is currently an assistant professor at the institute of electrical engineering, National Changhua University of Education, Taiwan, R.O.C. His research interests include digital signal processing, image processing, computer network, mobile computing and multimedia communications.

Authorized licensed use limited to: Emily Tannenbaum. Downloaded on August 11,2024 at 18:09:27 UTC from IEEE Xplore. Restrictions apply.